
















**METHODS AND COMPOSITIONS FOR ANALYZING NUCLEIC ACID  
MOLECULES UTILIZING SIZING TECHNIQUES****Publication number:** DE69814629 (T2)**Publication date:** 2004-03-25**Inventor(s):** VAN NESS JEFFREY [US]; TABONE C [US]; HOWBERT  
JEFFRY [US]; MULLIGAN T [US]**Applicant(s):** QIAGEN GENOMICS INC [US]**Classification:****- international:** *C12N15/09; C07F9/141; C07F9/24; C07F9/40; C07F9/547;  
C07H21/00; C08G69/08; C12Q1/68; G01N27/62; G01N33/53;  
C12N15/09; C07F9/00; C07H21/00; C08G69/00; C12Q1/68;  
G01N27/62; G01N33/53; (IPC1-7): C12Q1/68***- European:****Application number:** DE19986014629T 19980722**Priority number(s):** US19970898180 19970722; US19970898564 19970722;  
US19970898501 19970722; WO1998US15008 19980722**Also published as:** WO9905319 (A2)  
 WO9905319 (A3)  
 WO9905319 (A9)  
 PT990047 (E)  
 NZ501919 (A)  
 JP2009159985 (A)  
 JP2001511359 (T)  
 ES2200355 (T3)  
 EP0990047 (A2)  
 EP0990047 (B1)  
 DK0990047 (T3)  
 CN1270598 (A)  
 CA2297158 (A1)  
 AU738237 (B2)  
 AT240408 (T)

&lt;&lt; less

Abstract not available for DE 69814629 (T2)

Abstract of corresponding document: **WO 9905319 (A2)**

Tags and linkers specifically designed for a wide variety of nucleic acid reactions are disclosed, which are suitable for a wide variety of nucleic acid reactions wherein separation of nucleic acid molecules based upon size is required.

Data supplied from the **esp@cenet** database — Worldwide



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 698 14 629 T2** 2004.03.25

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 0 990 047 B1**

(51) Int Cl.<sup>7</sup>: **C12Q 1/68**

(21) Deutsches Aktenzeichen: **698 14 629.8**

(86) PCT-Aktenzeichen: **PCT/US98/15008**

(96) Europäisches Aktenzeichen: **98 936 928.5**

(87) PCT-Veröffentlichungs-Nr.: **WO 99/005319**

(86) PCT-Anmeldetag: **22.07.1998**

(87) Veröffentlichungstag

der PCT-Anmeldung: **04.02.1999**

(97) Erstveröffentlichung durch das EPA: **05.04.2000**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **14.05.2003**

(47) Veröffentlichungstag im Patentblatt: **25.03.2004**

(30) Unionspriorität:

**898180**                      **22.07.1997**                      **US**

**898564**                      **22.07.1997**                      **US**

**898501**                      **22.07.1997**                      **US**

(84) Benannte Vertragsstaaten:

**AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SE**

(72) Erfinder:

**VAN NESS, Jeffrey, Seattle, US; TABONE, C., John, Bothell, US; HOWBERT, Jeffry, Bellevue, US; MULLIGAN, T., John, Seattle, US**

(73) Patentinhaber:

**Qiagen Genomics, Inc., Bothell, Wash., US**

(74) Vertreter:

**BOEHMERT & BOEHMERT, 80336 München**

(54) Bezeichnung: **VERFAHREN UND VERBINDUNGEN ZUR ANALYSIERUNG VON NUKLEINSAEUREN DURCH MASSENSPEKTROMETRIE**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

**Beschreibung****TECHNISCHES GEBIET**

[0001] Die vorliegende Erfindung betrifft allgemein Verfahren und Zusammensetzungen zum Analysieren von Nukleinsäuremolekülen und, genauer gesagt, Markierungen, die in einer breiten Vielfalt von Nukleinsäurereaktionen eingesetzt werden können, bei denen die Abtrennung von Nukleinsäuremolekülen auf der Grundlage ihrer Größe erforderlich ist.

**HINTERGRUND DER ERFINDUNG**

[0002] Nachweis und Analyse von Nukleinsäuremolekülen sind mit die wichtigsten Techniken in der Biologie. Solche Techniken sind wesentlich in der Molekularbiologie und spielen eine sich schnell erweiternde Rolle im Rest der Biologie.

[0003] Im allgemeinen umfaßt eine Art von Analyse von Nukleinsäurereaktionen die Abtrennung von Nukleinsäuremolekülen auf der Grundlage ihrer Länge. Zum Beispiel ist eine im breiten Umfang verwendete Technik, die Polymerasekettenreaktion (PCR) (siehe U.S.-Patente Nrn. 4,683,195, 4,683,202 und 4,800,159), eine im breiten Umfang eingesetzte Technik geworden, um sowohl in einer Probe vorliegende Sequenzen zu identifizieren als auch DNA-Moleküle für weitere Manipulation zu synthetisieren.

[0004] Kurz gesagt, werden bei der PCR DNA-Sequenzen durch eine enzymatische Reaktion amplifiziert, die neue DNA-Stränge in entweder einer geometrischen oder linearen Art und Weise synthetisiert. Im Anschluß an die Amplifikation müssen die DNA-Sequenzen nachgewiesen und identifiziert werden. Wegen nicht-spezifischer Amplifikationen, die ansonsten die Analyse durcheinanderbringen würden, oder einer Reinheitsanforderung werden die PCR-Reaktionsprodukte im allgemeinen vor dem Nachweis einer Trennung unterzogen. Die Trennung auf der Grundlage der Größe (d. h. Länge) der Produkte liefert die nützlichste Information. Das Verfahren, das die höchste Auflösung von Nukleinsäuremolekülen ergibt, ist die elektrophoretische Trennung. Bei diesem Verfahren wird jede einzelne PCR-Reaktion auf ein geeignetes Gel aufgebracht und einem Spannungspotential ausgesetzt. Die Anzahl von Proben, die verarbeitet werden können, ist durch die Anzahl von Näpfen im Gel beschränkt. Auf den meisten Gelvorrichtungen können ungefähr 10 bis 64 Proben in einem einzigen Gel getrennt werden. Somit ist die Verarbeitung einer großen Anzahl von Proben sowohl arbeits- als auch materialintensiv.

[0005] Elektrophoretische Trennung muß mit irgendeinem Nachweissystem gekoppelt werden, um Daten zu erhalten. Nachweissysteme für Nukleinsäuren verwenden im allgemeinen, und fast ausschließlich, einen Einlagerungsfarbstoff oder eine radioaktive Markierung und weniger häufig eine nicht-radioaktive Markierung. Einlagerungsfarbstoffe, wie etwa Ethidiumbromid, sind einfach zu verwenden. Der Farbstoff wird während der Elektrophorese in die Gelmatrix einbezogen oder das Gel wird, im Anschluß an die Elektrophorese, in einer farbstoffhaltigen Lösung getränkt. Der Farbstoff kann in einigen Fällen direkt sichtbar gemacht werden, wird aber häufiger, und insbesondere für Ethidiumbromid, mit Licht (z. B. UV) angeregt, um zu fluoreszieren. Trotz dieser offensichtlich einfachen Anwendung haben solche Farbstoffe einige merkbare Nachteile. Erstens sind die Farbstoffe unempfindlich und es muß eine große Menge an Nukleinsäuremolekülen vorliegen, um die Produkte sichtbar zu machen. Zweitens sind die Farbstoffe typischerweise mutagen oder karzinogen.

[0006] Eine empfindlichere Nachweisteknik als Farbstoffe verwendet eine radioaktive (oder nicht-radioaktive) Markierung. Typischerweise wird entweder ein radioaktiv markiertes Nukleotid oder ein radioaktiv markierter Primer in die PCR-Reaktion einbezogen. Im Anschluß an die Trennung wird die radioaktive Markierung durch Autoradiographie "sichtbar gemacht". Obgleich empfindlicher leidet der Nachweis an Filmbeschränkungen, wie etwa Abweichung vom Reziprozitätsgesetz und Nicht-Linearität. Diese Beschränkungen können überwunden werden, indem die Markierung durch Phosphorbildanalyse nachgewiesen wird. Radioaktive Markierungen haben jedoch hohe Anforderungen an die Sicherheit, was die Ausnutzung von Ressourcen erhöht und spezialisierte Ausrüstung und Personaltraining erfordert. Aus solchen Gründen hat die Verwendung von nicht-radioaktiven Markierungen zunehmend an Popularität gewonnen. In solchen Systemen enthalten Nukleotide eine Markierung, wie etwa ein Fluorophor, Biotin oder Digoxin, die durch einen Antikörper oder ein anderes Molekül (z. B. anderes Teil eines Ligandenpaares) nachgewiesen werden kann, der (das) mit einem Enzym markiert ist, das mit einem chromogenen Substrat reaktiv ist. Diese Systeme haben nicht die Sicherheitsbedenken, wie oben beschrieben, verwenden aber Komponenten, die oft labil sind und nicht-spezifische Reaktionen liefern können, was zu einem hohen Hintergrund führt (d. h. niedrigem Signal-Rausch-Verhältnis).

[0007] US-A-5118605 betrifft Verfahren zum Testen eines Nukleinsäureanalyten. WO-A-9504160 betrifft ein Markierungsreagenz und ein Testverfahren.

## ZUSAMMENFASSUNG DER ERFINDUNG

[0008] Kurz gesagt stellt die vorliegende Erfindung Zusammensetzungen und Verfahren zur Verfügung, die in einer breiten Vielfalt von Ligandenpaarreaktionen eingesetzt werden können, bei denen die Trennung von interessierenden Molekülen, wie etwa Nukleinsäuremolekülen; auf der Grundlage ihrer Größe erforderlich ist. Repräsentative Beispiele für Verfahren, die mit dem Hintergrund der hierin zur Verfügung gestellten Offenbarung verbessert werden können, schließen PCR, Differential Display, RNA-Fingerprinting, PCR-SSCP, Oligo-Ligationsassays, Nukleaseverdauungsverfahren (z. B. Assays auf der Basis von Exo- und Endonuklease) und Didesoxy-Fingerprinting ein. Die hierin beschriebenen Verfahren können in einem weiten Bereich von Feldern eingesetzt werden, einschließlich z. B. bei der Entwicklung von Diagnostika für klinische Anwendung oder Forschung, der Bestimmung von Polymorphismen und der Entwicklung genetischer Karten.

[0009] In einem Aspekt der vorliegenden Erfindung wird eine Verbindung der Formel:



[0010] Zur Verfügung gestellt, worin  $T^{ms}$  eine organische Gruppe ist, die durch Massenspektrometrie nachweisbar ist, umfassend Kohlenstoff, mindestens eines von Wasserstoff und Fluor, und gegebenenfalls Atome, ausgewählt aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Jod;

[0011] L ist eine organische Gruppe, die es einer einmaligen  $T^{ms}$ -enthaltenden Gruppe ermöglicht, von dem Rest der Verbindung abgespalten zu werden, wobei die  $T^{ms}$ -enthaltende Gruppe eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterzogen wird, und die tertiäres Amin, quaternäres Amin oder organische Säure ist; und

[0012] X ist eine funktionelle Gruppe, ausgewählt aus Phosphoramidit und H-Phosphonat.

[0013] In einem anderen Aspekt stellt die vorliegende Erfindung ein Verfahren zur Bestimmung der Anwesenheit eines Single Nukleotide Polymorphismus in einem Nukleinsäureziel zur Verfügung, umfassend:

- a) Amplifizieren einer Sequenz eines Nukleinsäureziels, das einen Single Nukleotide Polymorphismus enthält;
- b) Erzeugen einer Einzelstrangform des Ziels;
- c) Kombinieren einer markierten Nukleinsäuresonde, die die Struktur  $T^{ms}\text{-L-X'}$  aufweist mit den amplifizierten Zielnukleinsäuremolekülen unter Bedingungen und für eine Zeitdauer, die ausreichend ist, um eine Hybridisierung der markierten Nukleinsäuresonde an komplementäre amplifizierte ausgewählte Zielnukleinsäuremoleküle zu ermöglichen, wobei die Markierung mit einem bestimmten Single Nukleotide Polymorphismus korrelierbar ist und durch Spektrometrie oder Potentiometrie nachweisbar ist;
- d) Abtrennen von nicht hybridisierter markierter Probe von hybridisierter markierter Probe durch eine Größenbestimmungs-Methodologie;
- e) Abspalten des Markers von der Sonde; und
- f) Nachweisen des Markers durch Spektrometrie oder Potentiometrie und Bestimmen der Anwesenheit des Single Nukleotide Polymorphismus.

[0014] In einem anderen Aspekt stellt die vorliegende Erfindung ein Verfahren zur Bestimmung der Anwesenheit eines Single Nukleotide Polymorphismus in einem Nukleinsäureziel zur Verfügung, umfassend:

- a) Amplifizieren einer Sequenz eines Nukleinsäureziels, das einen Single Nukleotide Polymorphismus enthält;
- b) Kombinieren eines markierten Nukleinsäureprimers, der die Struktur  $T^{ms}\text{-L-X'}$  aufweist, mit den amplifizierten Zielnukleinsäuremolekülen unter Bedingungen und für eine Zeitdauer, die ausreichend ist, um ein Annealing des markierten Nukleinsäureprimers an komplementäre amplifizierte ausgewählte Zielnukleinsäuremoleküle zu ermöglichen, wobei der Oligonukleotidprimer eine zu der Wildtyp-Sequenz oder dem Single Nukleotide Polymorphismus komplementäre äußerst 3'-liegende Base aufweist, wobei der Marker mit einem bestimmten Single Nukleotide Polymorphismus korrelierbar ist und durch Spektrometrie oder Potentiometrie nachweisbar ist;
- c) Verlängern des Primers, wobei ein zu dem Ziel komplementärer Strang synthetisiert wird, wenn die am weitesten 3'-liegende Base des Primers mit dem Ziel komplementär ist;
- d) Abtrennen von nicht verlängertem markierten Primer von verlängertem markierten Primer durch eine Größenbestimmungs-Methodologie;
- e) Abspalten des Markers von den Primern oder verlängerten Primern; und
- f) Nachweisen des Markers durch Spektrometrie oder Potentiometrie, und daraus Bestimmen der Anwesenheit des Single Nukleotide Polymorphismus.

[0015] In einem anderen Aspekt stellt die vorliegende Erfindung ein Verfahren zur Bestimmung der Menge eines spezifischen mRNA-Moleküls in einer Nukleinsäurepopulation zur Verfügung, umfassend:

- a) Überführen einer RNA-Population in eine cDNA-Population;
- b) Hinzufügen einer einzelsträngigen Nukleinsäure (interner Standard), die eine Vielzahl von Single Nukleotide Polymorphismen enthält, die ansonsten identisch zu dem cDNA-Ziel ist;



- c) Amplifizieren einer spezifischen Sequenz des cDNA-Ziels;
- d) Coamplifizieren des internen Standards, wobei der interne Standard dieselbe Länge aufweist, wie das cDNA-Amplicon;
- e) Erzeugen einer einzelsträngigen Form des Ziels;
- f) Kombinieren eines Sets von markierten Nukleinsäuresonden, die die Struktur  $T^{ms}$ -L-X' aufweisen, mit der amplifizierten Ziel-cDNA und dem amplifizierten internen Standard unter Bedingungen für eine Zeitdauer, die ausreichend ist, um eine Hybridisierung der markierten Nukleinsäuresonde an komplementäre ausgewählte Ziel-cDNA und interne Standardsequenzen zu erlauben, wobei der Marker mit einer bestimmten cDNA-Sequenz korrelierbar ist und ein zweiter Marker ist mit dem internen Standard korrelierbar, und ist durch Spektrometrie oder Potentiometrie nachweisbar;
- g) Abtrennen von nicht hybridisierter markierter Sonde von hybridisierter markierter Sonde durch eine Größenbestimmungs-Methodologie;
- h) Abspalten des Markers von den Sonden;
- i) Nachweisen der Marker durch Spektrometrie oder Potentiometrie; und
- j) Nehmen des Verhältnisses von Marker, der zu cDNA korreliert ist, gegenüber Marker, der mit dem internen Standard korreliert ist und daraus bestimmen der Menge der cDNA, wodurch die Menge der spezifischen mRNA in einer Nukleinsäurepopulation bestimmt wird.

[0016] In den oben genannten Verfahren weist die markierte Nukleinsäure die Struktur  $T^{ms}$ -L-X' auf, wobei X' die Nukleinsäure ist und  $T^{ms}$  und L wie oben definiert sind.

[0017] Innerhalb eines Aspekts der vorliegenden Erfindung werden Verfahren zum Bestimmen der Identität eines Nukleinsäuremoleküls zur Verfügung gestellt, die die Schritte umfassen von (a) Erzeugen markierter Nukleinsäuremoleküle aus einem oder mehreren ausgewählten Target-Nukleinsäuremolekülen, wobei eine Markierung mit einem bestimmten Nukleinsäurefragment korrelativ und durch Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie nachweisbar ist, (b) Abtrennen der markierten Fragmente nach Größe, (c) Abspalten der Markierungen von den markierten Fragmenten und (d) Nachweisen der Markierungen durch Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie und Bestimmen der Identität der Nukleinsäuremoleküle daraus.

[0018] In einem verwandten Aspekt der Erfindung werden Verfahren zum Nachweisen eines ausgewählten Nukleinsäuremoleküls zur Verfügung gestellt, die die Schritte umfassen: (a) Kombinieren markierter Nukleinsäuresonden mit Target-Nukleinsäuremolekülen unter Bedingungen und für einen Zeitraum, die ausreichend sind, um Hybridisierung einer markierten Nukleinsäuresonde an eine komplementäre ausgewählte Target-Nukleinsäuresequenz zu ermöglichen, wobei eine markierte Nukleinsäuresonde durch Nicht-Fluoreszenzspektrometrie oder Potentiometrie nachweisbar ist, (b) Verändern der Größe hybridisierter markierter Sonden, nicht-hybridisierter Sonden oder Target-Moleküle oder der Sonden:Target-Hybride, (c) Abtrennen der markierten Sonden nach Größe, (d) Abspalten der Markierungen von den markierten Sonden und (e) Nachweisen der Markierungen durch Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie und Nachweisen des ausgewählten Nukleinsäuremoleküls daraus.

[0019] In weiteren Aspekten werden Verfahren zum Genotypisieren eines ausgewählten Organismus zur Verfügung gestellt, die die Schritte umfassen: (a) Erzeugen markierter Nukleinsäuremoleküle aus einem ausgewählten Target-Molekül, wobei eine Markierung mit einem bestimmten Fragment korrelativ ist und durch Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie nachgewiesen werden kann, (b) Abtrennen der markierten Moleküle nach Sequenzlänge, (c) Abspalten der Markierung vom markierten Molekül und (d) Nachweisen der Markierung durch Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie und Bestimmen des Genotyps des Organismus daraus.

[0020] In einem weiteren Aspekt werden Verfahren zum Genotypisieren eines ausgewählten Organismus zur Verfügung gestellt, die die Schritte umfassen: (a) Kombinieren eines markierten Nukleinsäuremoleküls mit einem ausgewählten Target-Molekül unter Bedingungen und für einen Zeitraum, die ausreichend sind, um Hybridisierung des markierten Moleküls an das Target-Molekül zu ermöglichen, wobei eine Markierung mit einem bestimmten Fragment korrelativ ist und durch Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie nachgewiesen werden kann, (b) Abtrennen der markierten Fragmente nach Sequenzlänge, (c) Abspalten der Markierung vom markierten Fragment und (d) Nachweisen der Markierung durch Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie und Bestimmen des Genotyps des Organismus daraus.

[0021] Im Kontext der vorliegenden Erfindung sollte "biologische Proben" so verstanden werden, daß diese nicht nur Proben einschließen, die von lebenden Organismen erhalten werden (z. B. Säugern, Fischen, Bakterien, Parasiten, Viren, Pilzen und dergleichen) oder aus der Umwelt (z. B. Luft, Wasser oder Feststoffproben), sondern auch biologische Materialien, die künstlich oder synthetisch hergestellt sein können (z. B. Phagenbibliotheken, Bibliotheken organischer Moleküle, Pools genomischer Klone, cDNA-Klone, RNA-Klone oder dergleichen). Repräsentative Beispiele für biologische Proben schließen biologische Flüssigkeiten (z. B. Blut, Samen, Hirn-Rückenmarks-Flüssigkeit, Urin), biologische Zellen (z. B. Stammzellen, B- oder T-Zellen, Leberzellen, Fibroblasten oder dergleichen) und biologische Gewebe ein. Schließlich schließen repräsentative Beispiele

le für Organismen, die genotypisiert werden können, praktisch jeden einzelligen oder mehrzelligen Organismus ein, wie etwa Warmblüter, Säuger oder Wirbeltiere (z. B. Menschen, Schimpansen, Makaken, Pferde, Kühe, Schweine, Schafe, Hunde, Katzen, Ratten und Mäuse sowie Zellen von irgendeinem dieser), Bakterien, Parasiten, Viren, Pilze und Pflanzen.

[0022] In verschiedenen Ausführungsformen der oben beschriebenen Verfahren können die Nukleinsäuresonden und/oder -moleküle der vorliegenden Erfindung durch z. B. einen Ligations-, Spaltungs- oder Extensions(z. B. PCR)-Reaktion erzeugt werden. In weiteren verwandten Aspekten können die Nukleinsäuresonden oder -moleküle mit nicht-3'-markierten Oligonukleotid-Primern (z. B. 5'-markierten Oligonukleotid-Primern) oder Didesoxynukleotidterminatoren markiert sein.

[0023] In weiteren Ausführungsformen der Erfindung können 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 250, 300, 350, 400, 450 oder mehr als 500 unterschiedliche und einzigartig markierte Moleküle in einer gegebenen Reaktion gleichzeitig eingesetzt werden, wobei jede Markierung für ein ausgewähltes Nukleinsäuremolekül oder -fragment, oder -sonde, einzigartig ist und separat identifiziert werden kann.

[0024] Die Markierungen werden durch Massenspektrometrie nachgewiesen. Repräsentative Beispiele für schließen time-of-flight-Massenspektrometrie, Quadrapol-Massenspektrometrie, Magnetsektor-Massenspektrometrie und Elektrosektor-Massenspektrometrie ein. Spezifische Ausführungsformen solcher Techniken schließen Ionenfallen-Massenspektrometrie, Elektrosprayionisations-Massenspektrometrie, Ionenspray-Massenspektrometrie, Flüssigionisations-Massenspektrometrie, Atmosphärendruckionisations-Massenspektrometrie, Elektronenionisations-Massenspektrometrie, Massenspektrometrie mit Ionisation durch Beschuß mit schnellen Atomen, MALDI-Massenspektrometrie, Photoionisations-Flugzeit-Massenspektrometrie, Laser-Tröpfchen-Massenspektrometrie, MALDI-TOF-Massenspektrometrie, APCI-Massenspektrometrie, Nanospray-Massenspektrometrie, Massenspektrometrie mit Ionisation zerstäubten Sprays, Massenspektrometrie mit chemischer Ionisation, Resonanzionisations-Massenspektrometrie, Sekundärionisations-Massenspektrometrie und Thermospray-Massenspektrometrie ein.

[0025] In noch weiteren Ausführungsformen der Erfindung können die Target-Moleküle, hybridisierten markierten Sonden, nicht-hybridisierten Sonden oder Target-Moleküle, Sonden:Target-Hybride oder markierten Nukleinsäuresonden oder -moleküle von anderen Molekülen unter Verwendung von Verfahren abgetrennt werden, die zwischen der Größe von Molekülen unterscheiden (entweder tatsächliche lineare Größe oder dreidimensionale Größe). Repräsentative Beispiele solcher Verfahren schließen Gelelektrophorese, Kapillarelektrophorese, Mikrokanalelektrophorese, HPLC, Ausschußchromatographie, Filtration, Polyacrylamidgelelektrophorese, Flüssigchromatographie, Umkehrausschußchromatographie, Ionenaustauschchromatographie, Ionenaustauschchromatographie, Pulsfeldelektrophorese, Feldinversionselektrophorese, Dialyse und Fluoreszenz-aktivierte Flüssigtröpfchensortierung. Alternativ können die Target-Moleküle, hybridisierten markierten Sonden, nicht-hybridisierten Sonden oder Target-Moleküle, Sonden:Target-Hybride oder markierten Nukleinsäuresonden oder -moleküle an einen festen Träger (z. B. Hohlfasern (Amicon Corporation, Danvers, Mass.), Perlen (Polysciences, Warrington, Pa.), Magnetperlen (Robbin Scientific, Mountain View, Calif.), Platten, Schalen und Kolben (Corning Glass Works, Corning, N.Y.), Maschengitter (Becton Dickinson, Mountain View, Calif.), Siebe und feste Fasern (siehe Edelman et al., U.S.-Patent Nr. 3,843,324; siehe auch Kuroda et al., U.S.-Patent Nr. 4,416,777), Membranen (Millipore Corp., Bedford, Mass.) und Tauschstäbchen) gebunden sein. Wenn das erste oder zweite Teil oder freiliegende Nukleinsäuren an einen festen Träger sind, können die hierin offenbarten Verfahren in bestimmten Ausführungsformen der Erfindung außerdem den Schritt umfassen, daß nicht-gebundenes Material vom Träger abgewaschen wird.

[0026] In weiteren Ausführungsformen können die markierten Nukleinsäuremoleküle oder -sonden mit solchen Verfahren gespalten werden, wie chemischen, Oxidations-, Reduktions-, säurelabilen, basenlabilen, enzymatischen, elektrochemischen, wärme- und photolabilen Verfahren. In weiteren Ausführungsformen können die Schritte Abtrennen, Spalten und Nachweisen in einer kontinuierlichen Art und Weise durchgeführt werden, z. B. auf einer einzigen Vorrichtung, die automatisiert sein kann.

[0027] In bestimmten Ausführungsformen der Erfindung werden die Größe der hybridisierten markierten Sonden, nicht-hybridisierten Sonden oder Targetmoleküle oder Sonden:Target-Hybride mit einem Verfahren verändert, das ausgewählt ist aus der Gruppe, die aus Polymerase-Extension, Ligation, Exonuklease-Verdau, Endonuklease-Verdau, Restriktionsenzym-Verdau, ortsspezifischem Rekombinase-Verdau, Ligation, Mismatch-Verdau mit spezifischer Nuklease, Methylierungs-spezifischer Nuklease-Verdau, kovalente Bindung von Sonde an Target und Hybridisierung besteht.

[0028] Die Verfahren und Zusammensetzungen, die hierin beschrieben sind, können in einer breiten Vielfalt von Anwendungen eingesetzt werden, einschließlich z. B. Identifizieren von PCR-Amplikons, RNA-Fingerprinting, Differentialdisplay, Einzelstrangkonformationspolymorphismusnachweis, Didesoxy-Fingerprinting, Restriktionskarten- und Restriktionsfragmentlängenpolymorphismen, DNA-Fingerprinting, Genotypisierung, Mutationsnachweis, Oligonukleotidligationsassay, Sequenz-spezifische Amplifikationen, für Diagnose, Forensik, Identifikation, Entwicklungsbiologie, Biologie, Molekularmedizin, Toxikologie, Tierzüchtung.

[0029] Diese und weitere Aspekte der vorliegenden Erfindung werden bei Bezugnahme auf die folgende de-

taillierte Beschreibung und die beigelegten Zeichnungen deutlich werden.

#### KURZE BESCHREIBUNG DER ZEICHNUNGEN

[0030] **Fig. 1** zeigt das Flußdiagramm für die Synthese von Pentafluorphenylestern chemisch spaltbarer Massenspektroskopiemarkierungen, um Markierungen mit Carbonsäureamid-Endgruppen freizusetzen.

[0031] **Fig. 2** zeigt das Flußdiagramm für die Synthese von Pentafluorphenylestern chemisch spaltbarer Massenspektroskopiemarkierungen, um Markierungen mit Carbonsäure-Endgruppen freizusetzen.

[0032] **Fig. 3-6 und 8** zeigen das Flußdiagramm für die Synthese von Tetrafluorphenylestern eines Satzes von 36 photochemisch spaltbaren Massenspektroskopiemarkierungen.

[0033] **Fig. 7** zeigt das Flußdiagramm für die Synthese eines Satzes von 36 photochemisch spaltbaren Massenspektroskopiemarkierungen mit Amin-Endgruppen.

[0034] **Fig. 9** zeigt die Synthese von 36 photochemisch spaltbaren Massenspektroskopie-markierten Oligonukleotiden, hergestellt aus dem entsprechenden Satz von 36 Tetrafluorphenylestern photochemisch spaltbarer Massenspektroskopiemarkierungssäuren.

[0035] **Fig. 10** zeigt die Synthese von 36 photochemisch spaltbaren Massenspektroskopie-markierten Oligonukleotiden, hergestellt aus dem entsprechenden Satz von 36 photochemisch spaltbaren Massenspektroskopiemarkierungen mit Amin-Endgruppen.

[0036] **Fig. 11** veranschaulicht den gleichzeitigen Nachweis von mehreren Markierungen durch Massenspektrometrie.

[0037] **Fig. 12** zeigt das Massenspektrogramm der Alpha-Cyano-Matrix allein.

[0038] **Fig. 13** zeigt ein modular konstruiertes markiertes Nukleinsäurefragment.

[0039] **Fig. 14A-Fig. 14I** zeigen die Abtrennung von DNA-Fragmenten durch HPLC unter der Verwendung einer Vielzahl von verschiedenen Pufferlösungen.

[0040] **Fig. 15** zeigt eine schematische Darstellung von genetischem Fingerprinting- und Differential-Display-Systemen in Übereinstimmung mit einer beispielhaften Ausführungsform der vorliegenden Erfindung.

[0041] **Fig. 16** zeigt eine schematische Darstellung von genetischem Fingerprinting- und Differential-Display-Systemen in Übereinstimmung mit einer beispielhaften Ausführungsform der vorliegenden Erfindung.

[0042] **Fig. 17** zeigt eine schematische Darstellung von Testsystemen in Übereinstimmung mit einer exemplarischen Ausführungsform der vorliegenden Erfindung.

[0043] **Fig. 18** ist eine schematische Darstellung von Assay-Systemen in Übereinstimmung mit einer exemplarischen Ausführungsform der vorliegenden Erfindung.

[0044] **Fig. 19A und 19B** zeigen die Herstellung eines absaltbaren Markers der vorliegenden Erfindung.

[0045] **Fig. 20A und 20B** illustrieren die Herstellung eines absaltbaren Markers der vorliegenden Erfindung.

[0046] **Fig. 21** zeigt die Herstellung einer intermediären Verbindung; die bei der Herstellung eines spaltbaren Markers der Erfindung brauchbar ist.

[0047] **Fig. 22A, 22B und 22C** zeigen die Synthese-Methodologie für die Herstellung eines photo-absaltbaren Massenspektrometrie-nachweisbaren Markers gemäß der vorliegenden Erfindung.

[0048] **Fig. 23** zeigt die Ergebnisse aus einem Assay, der die Genexpression mit CMST-markierten ODNs verfolgte.

[0049] **Fig. 24-28** zeigen Phosphoramidit-Chemie, die in einem Beispiel hier vollständig beschrieben wird.

#### DETAILLIERTE BESCHREIBUNG DER ERFINDUNG

[0050] Wie oben angegeben, stellt die vorliegende Erfindung Zusammensetzungen und Verfahren zum Analysieren von Nukleinsäuremolekülen zur Verfügung, bei denen die Trennung von Nukleinsäuremolekülen auf der Grundlage ihrer Größe erforderlich ist. Die vorliegenden Verfahren ermöglichen den gleichzeitigen Nachweis von interessierenden Molekülen, die Nukleinsäuren und Fragmente, Proteine, Peptide, usw. einschließen.

[0051] Die vorliegende Erfindung stellt eine neue Klasse von Markern für genomische Messungen zur Verfügung, die eine Assay-Plattform zur Verfügung stellen, die mit dem Ausmaß von Messungen kompatibel ist, die für die Analyse von komplexen Genomen erforderlich ist. Diese neue Markertechnologie ist aus Massenspektrometrie, Markern, die mit einem Standard-Quadrupol-Massenspektrometerdetektor (MSD) unter der Verwendung von atmosphärischem Druck chemischer Ionisierung (positiver Modus) nachgewiesen werden, zusammengesetzt. Die Technologieplattform verwendet einen MSD zum Nachweis der bekannten Molekulargewichts-Massenspektrometermarkern, ganz ähnlich einem Dioden-Array-Detektor. Die Marker können durch kombinatorische Chemie-Ansätze unter der Verwendung eines primären Gerüsts synthetisiert werden, auf das spezifische Masse-Einsteller angefügt sind. Die Marker sind so aufgebaut, um reversibel an Oligonukleotide angebracht zu sein, die entweder als Primer im PCR-Ansatz verwendet werden können oder als Sonden in Hybridisierungsassays verwendet werden. Bei der Vervollständigung jeder Zahl von Assay-Schritten wird der Marker/die Sonde oder Marker/Primer einer Abspaltungsreaktion unterzogen, bevorzugterweise Photoabspaltung.

tung, und wenn die Marker-Massenspektrometrie-nachweisbar sind, werden die Marker durch APCI ionisiert und die Massenidentität des Markers wird durch Massenspektrometrie nachgewiesen. Die Marker können dazu verwendet werden, um die Identität einer Sequenz und die Probenidentifizierung zu kartieren.

[0052] Kurz gesagt, stellt die vorliegende Erfindung in einem Aspekt Verbindungen zur Verfügung, bei denen ein interessierendes Molekül, oder eine Vorstufe dazu, über eine labile Bindung (oder labile Bindungen) an eine Markierung gebunden ist. So können Verbindungen der Erfindung so betrachtet werden, daß sie die allgemeine Formel:



[0053] Aufweisen, in der T die Markierungskomponente ist, L die Linkerkomponente ist, die entweder eine labile Bindung ist oder enthält, und X entweder die Komponente "Interessierendes Molekül" (MOI) oder eine Komponente mit einer funktionellen Gruppe ( $L_n$ ), durch die das MOI mit T-L verknüpft werden kann, ist. Verbindungen der Erfindung können daher durch die spezifischeren allgemeinen Formeln dargestellt werden:



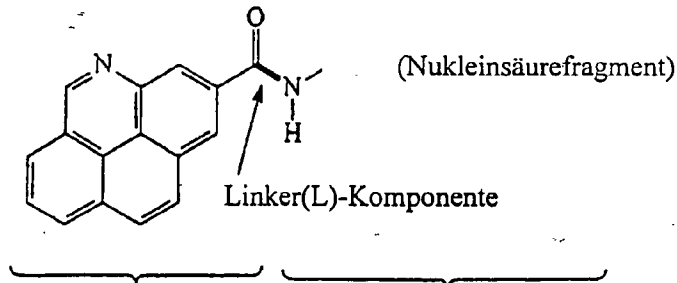
und



[0054] Aus unten genauer beschriebenen Gründen können Sätze von T-L-MOI-Verbindungen gezielt Bedingungen unterworfen werden, die bewirken, daß die labile(n) Bindungen) bricht (brechen), wodurch eine Markierungseinheit vom Rest der Verbindung freigesetzt wird. Die Markierungseinheit wird dann durch eine oder mehrere analytische Techniken charakterisiert, um dadurch direkte Information über die Struktur der Markierungseinheit zu liefern und (am wichtigsten) indirekte Information über die Identität des entsprechenden MOI.

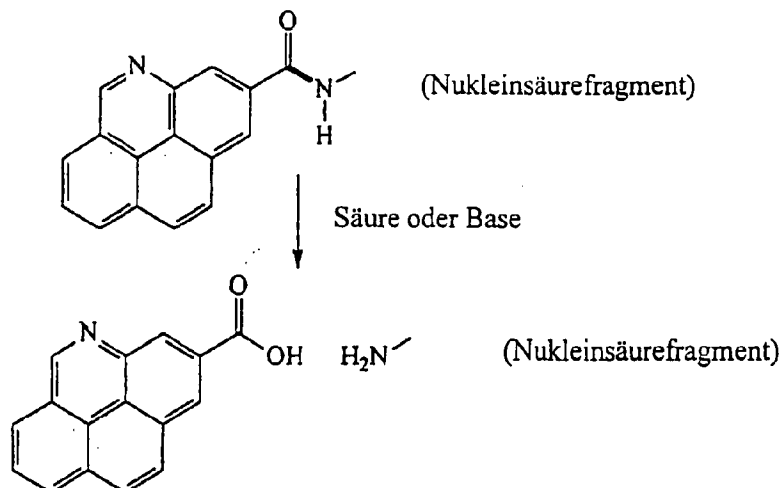
[0055] Als ein einfaches veranschaulichendes Beispiel für eine repräsentative Verbindung der Erfindung, bei der L eine direkte Bindung ist, wird Bezug genommen auf die folgende Struktur (i):

Struktur (i)



[0056] In Struktur (i) ist T eine stickstoffhaltige polycyclische aromatische Einheit, die an eine Carbonylgruppe gebunden ist, X ist ein MOI (und spezifisch ein Nukleinsäurefragment, das in einer Aminogruppe endet) und L ist die Bindung, die eine Amidgruppe bildet. Die Amidbindung ist relativ zu den Bindungen in T labil, weil eine Amidbindung, wie im Stand der Technik anerkannt, durch saure oder basische Bedingungen chemisch gespalten (gebrochen) werden kann, die die Bindungen in der Markierungskomponente unverändert lassen. So kann eine Markierungseinheit (d. h. das Spaltprodukt, das T enthält) freigesetzt werden, wie unten dargestellt:

Struktur (i)

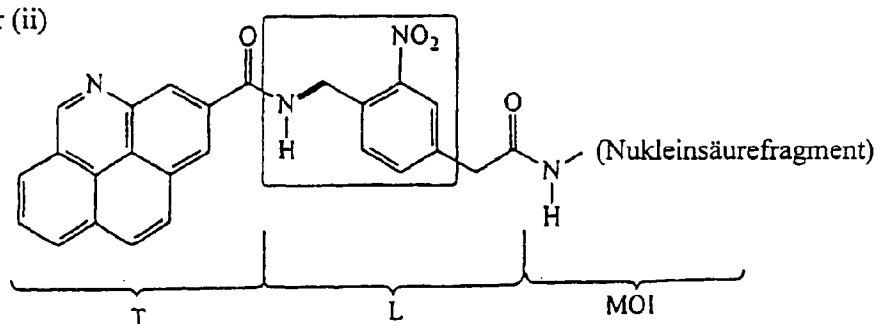


Markierungseinheit

Rest der Verbindung

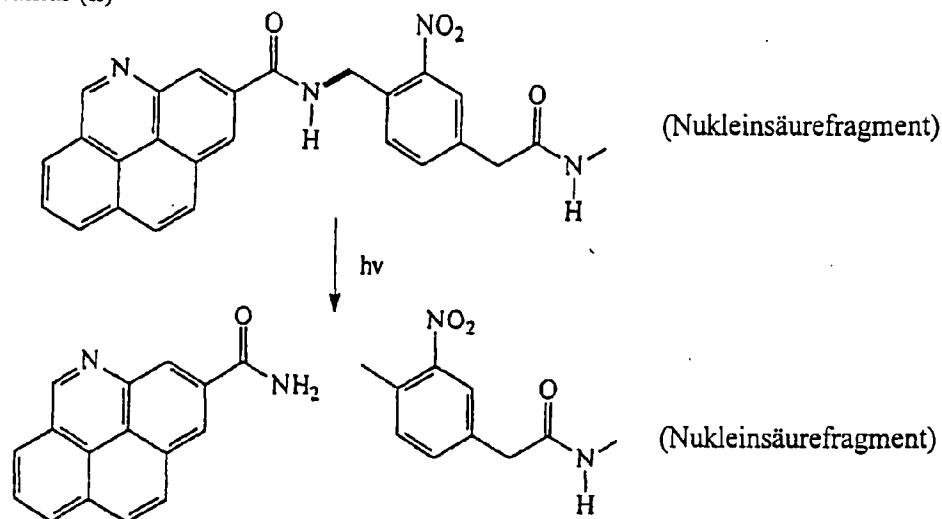
[0057] Der Linker L kann jedoch mehr sein als lediglich eine direkte Bindung, wie in dem folgenden veranschaulichenden Beispiel gezeigt, in dem Bezug genommen wird auf eine weitere repräsentative Verbindung der Erfindung der unten dargestellten Struktur (ii):

Struktur (ii)



[0058] Es ist gut bekannt, daß Verbindungen mit einer ortho-Nitrobenzylamin-Einheit (siehe eingerahmte Atome in Struktur (ii)) photolytisch insofern instabil sind, als die Einwirkung aktinischer Strahlung einer spezifizierten Wellenlänge auf solche Verbindungen eine selektive Spaltung der Benzylamin-Bindung (siehe Bindung, die in Struktur (ii) mit einer dicken Linie bezeichnet ist) bewirkt. So hat Struktur (ii) dieselben T- und MOI-Gruppen wie Struktur (i) die Linker-Gruppe enthält jedoch mehrere Atome und Bindungen, innerhalb derer sich eine besonders labile Bindung befindet. Photoyse von Struktur (ii) setzt somit eine Markierungseinheit (T-enthaltende Einheit) vom Rest der Verbindung frei, wie unten dargestellt.

Struktur (ii)



Markierungseinheit

Rest der Komponente

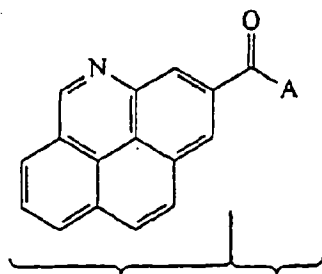
[0059] Die Erfindung stellt somit Verbindungen zur Verfügung, die, bei Einwirkung geeigneter Spaltungsbedingungen, eine Spaltungsreaktion durchlaufen, so daß eine Markierungseinheit vom Rest der Verbindung freigesetzt wird. Verbindungen der Erfindung können als die Markierungseinheit, das MOI (oder eine Vorstufe dazu,  $L_n$ ) und die labile(n) Bindung(en), die die zwei Gruppen miteinander verbindet (verbinden), beschrieben werden. Alternativ können die Verbindungen der Erfindung als die Komponenten beschrieben werden, aus denen sie gebildet sind. So können die Verbindungen wie folgt als das Reaktionsprodukt eines Markierungs-Reaktanten, eines Linker-Reaktanten und eines MOI-Reaktanten beschrieben werden.

[0060] Der Markierungs-Reaktant besteht aus einem chemischen Griff ( $T_n$ ) und einer variablen Komponente ( $T_{vc}$ ), so daß der Markierungs-Reaktant so betrachtet wird, daß er die allgemeine Struktur hat:



[0061] Um diese Nomenklatur zu veranschaulichen, kann Bezug genommen werden auf Struktur (iii), die einen Markierungs-Reaktanten zeigt, der verwendet werden kann, um die Verbindung von Struktur (ii) herzustellen. Der Markierungs-Reaktant mit Struktur (iii) enthält eine variable Markierungskomponente und einen Markierungsgriff, wie unten dargestellt:

Struktur (iii)

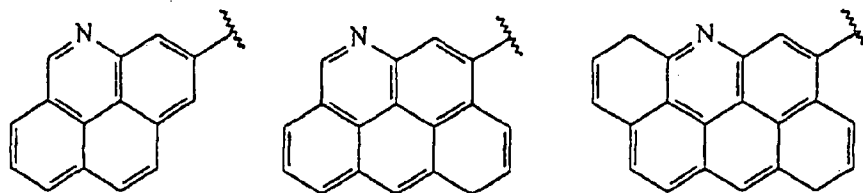


Variable Markierungskomponente      Markierungsgriff

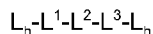
[0062] In Struktur (iii) liefert der Markierungsgriff ( $-C(=O)-A$ ) nur einen Weg für die Reaktion des Markierungs-Reaktanten mit dem Linker-Reaktanten, um eine T-L-Einheit zu bilden. Die Gruppe "A" in Struktur (iii) zeigt, daß die Carboxylgruppe sich in einem chemisch aktiven Zustand befindet, so daß sie für Kopplung mit anderen Griffen bereitsteht. "A" kann z. B. eine Hydroxylgruppe oder Pentafluorphenoxy sein, unter vielen anderen Möglichkeiten. Die Erfindung liefert eine große Anzahl möglicher Markierungsgriffe, die an eine variable

Markierungskomponente gebunden werden können, wie im Detail unten diskutiert. Die variable Markierungskomponente ist somit ein Teil von "T" in der Formel T-L-X und wird auch Teil der Markierungseinheit sein, die sich aus der Reaktion bildet, die L spaltet.

[0063] Wie ebenfalls unten genauer diskutiert, wird die variable Markierungskomponente so genannt, weil es bei der Herstellung von Sätzen von Verbindungen gemäß der Erfindung erwünscht ist, daß Mitglieder eines Satzes einzigartige variable Komponenten aufweisen, so daß die einzelnen Mitglieder voneinander mit einer analytischen Technik unterschieden werden können. Als ein Beispiel kann die variable Markierungskomponente von Struktur (iii) Mitglied des folgenden Satzes sein, in dem die Mitglieder des Satzes durch ihre UV- oder Massenspektren unterschieden werden können:

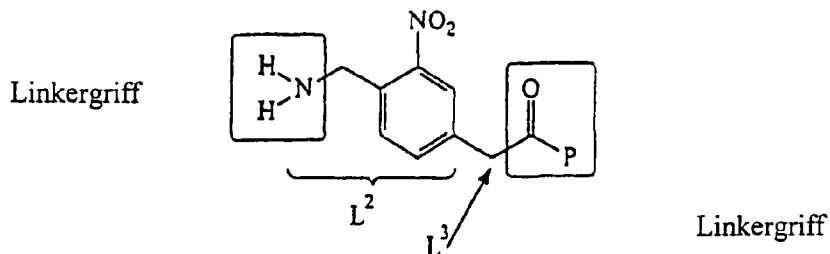


[0064] In ähnlicher Weise kann der Linker-Reaktant anhand seiner chemischen Griffe (es gibt notwendigerweise wenigstens zwei, von denen jeder als  $L_n$  bezeichnet werden kann), beschrieben werden, die eine labile Linkerkomponente flankieren, wobei die labile Linkerkomponente aus der erforderlichen labilen Einheit ( $L^2$ ) und fakultativen labilen Einheiten ( $L^1$  und  $L^3$ ) besteht, wobei die fakultativen labilen Einheiten effektiv dazu dienen,  $L^2$  von den Griffen  $L_n$  zu trennen, und die erforderliche labile Einheit dazu dient, eine labile Bindung innerhalb der labilen Linkerkomponente bereitzustellen. So kann der Linker-Reaktant so betrachtet werden, daß er die allgemeine Formel hat:



[0065] Die Nomenklatur, die verwendet wird, um den Linker-Reaktanten zu beschreiben, kann im Hinblick auf Struktur (iv) veranschaulicht werden, die sich wieder von der Verbindung von Struktur (ii) ableitet:

Struktur (iv)



[0066] Wie Struktur (iv) veranschaulicht, können Atome in mehr als eine funktionelle Rolle erfüllen. So dient, in Struktur (iv), der Benzylstickstoff als ein chemischer Griff, um zu ermöglichen, daß der Linker-Reaktant sich an den Markierungs-Reaktanten über eine Amidbildungsreaktion bindet, und dient anschließend auch insofern als ein notwendiger Teil der Struktur der labilen Einheit  $L^2$ , als die Benzyl-Kohlenstoff-Stickstoff-Bindung besonders anfällig ist für photolytische Spaltung. Struktur (iv) veranschaulicht auch, daß ein Linker-Reaktant eine  $L^3$ -Gruppe aufweisen kann (in diesem Fall eine Methylengruppe), obgleich er keine  $L^1$ -Gruppe hat. In ähnlicher Weise können Linker-Reaktanten eine  $L^1$ -Gruppe aufweisen, aber keine  $L^3$ -Gruppe, oder können  $L^1$ - und  $L^3$ -Gruppen aufweisen oder können weder  $L^1$ - noch  $L^3$ -Gruppen aufweisen. In Struktur (iv) zeigt das Vorhandensein der Gruppe "P" neben der Carbonylgruppe, daß die Carbonylgruppe vor der Reaktion geschützt ist. Bei Vorgabe dieser Konfiguration kann die aktivierte Carboxylgruppe des Markierungs-Reaktanten (iii) sauber mit der Aminogruppe des Linker-Reaktanten (iv) reagieren, um eine Amidbindung zu bilden und eine Verbindung der Formel T-L- $L_n$  zu ergeben.

[0067] Der MOI-Reaktant ist eine in geeigneter Weise reaktive Form eines interessierenden Moleküls. Wenn das interessierende Molekül ein Nukleinsäurefragment ist, ist ein geeigneter MOI-Reaktant ein Nukleinsäurefragment, das durch seine 5'-Hydroxylgruppe an eine Phosphodiestergruppe gebunden ist und anschließend an eine Alkyl-Kette, die in einer Aminogruppe endet. Diese Aminogruppe kann dann mit der Carbonylgruppe von Struktur (iv) reagieren (natürlich nach Entfernen der Schutzgruppe der Carbonylgruppe und vorzugsweise nach anschließender Aktivierung der Carbonylgruppe für die Reaktion mit der Aminogruppe), um dadurch das MOI an den Linker zu binden.

[0068] Wenn sie in einer chronologischen Reihenfolge betrachtet wird, kann die Erfindung so gesehen werden, daß sie einen Markierungs-Reaktanten (mit einem chemischen Markierungsgriff und einer variablen Markierungskomponente), einen Linker-Reaktanten (mit zwei chemischen Linkergriffen, einer erforderlichen labilen Einheit und 0–2 fakultativen labilen Einheiten) und einen MOI-Reaktanten (mit einer Komponente "Interessierendes Molekül" und einem chemischen Griff des interessierenden Moleküls) nimmt, um T-L-MOI zu bilden. So werden, um T-L-MOI zu bilden, entweder der Markierungs-Reaktant und der Linker-Reaktant zunächst miteinander umgesetzt, um T-L-L<sub>n</sub> zu liefern, und anschließend wird der MOI-Reaktant mit T-L-L<sub>n</sub> umgesetzt, um T-L-MOI zu liefern, oder (weniger bevorzugt) werden der Linker-Reaktant und der MOI-Reaktant zunächst miteinander umgesetzt, um Lh-L-MOI zu liefern, und anschließend wird Lh-L-MOI mit dem Markierungs-Reaktanten umgesetzt, um T-L-MOI zu liefern. Aus Gründen der Bequemlichkeit werden Verbindungen mit der Formel T-L-MOI als der Markierungs-Reaktant, der Linker-Reaktant und der MOI-Reaktant beschrieben werden, die verwendet werden können, um solche Verbindungen zu bilden. Dieselben Verbindungen von Formel T-L-MOI könnten natürlich mit anderen (typischerweise aufwendigeren) Verfahren hergestellt werden und würden immer noch unter den Schutzzumfang der erfinderschen T-L-MOI-Verbindungen fallen.

[0069] In jedem Falle sorgt die Erfindung dafür, daß eine T-L-MOI-Verbindung Spaltungsbedingungen unterworfen wird, so daß eine Markierungseinheit vom Rest der Verbindung freigesetzt wird. Die Markierungseinheit wird wenigstens die variable Markierungskomponente umfassen und wird typischerweise zusätzlich einige oder alle Atome vom Markierungsgriff, einige oder alle Atome vom Linkergriff, die verwendet wurden, um das Markierungs-Reaktant an den Linker-Reaktanten zu binden, die fakultative labile Einheit L<sup>1</sup>, wenn diese Gruppe in T-L-MOI vorhanden war, umfassen und wird vielleicht einen Teil der erforderlichen labilen Einheit L<sup>2</sup> enthalten, in Abhängigkeit von der genauen Struktur von L<sup>2</sup> und der Natur der Spaltungschemie. Aus Gründen der Bequemlichkeit kann die Markierungseinheit als die T-enthaltende Einheit bezeichnet werden, weil T typischerweise den Hauptteil (im Hinblick auf die Masse) der Markierungseinheit ausmachen wird.

[0070] Nach dieser Einführung in einen Aspekt der vorliegenden Erfindung werden die verschiedenen Komponenten T, L und X im Detail beschrieben werden. Diese Beschreibung beginnt mit den folgenden Definitionen bestimmter Begriffe, die im weiteren bei der Beschreibung von T, L und X verwendet werden.

[0071] Wie hierin verwendet, bedeutet der Begriff "Nukleinsäurefragment" ein Molekül, das komplementär zu einem ausgewählten Target-Nukleinsäuremolekül ist (d. h. komplementär zu dem gesamten Molekül oder einem Teil desselben) und aus der Natur gewonnen oder synthetisch oder rekombinant hergestellt werden kann, einschließlich nicht natürlich vorkommender Moleküle, und kann je nach Eignung in doppel- oder einzelsträngiger Form vorliegen; und schließt ein Oligonukleotid (z. B. DNA oder RNA), einen Primer, eine Sonde, ein Nukleinsäureanalog (z. B. PNA), ein Oligonukleotid, das in einer 5'-nach-3'-Richtung mit einer Polymerase verlängert ist, eine Nukleinsäure, die chemisch oder enzymatisch gespalten ist, eine Nukleinsäure, die mit einem Didesoxyterminator endet oder am 3'- oder 5'-Ende mit einer Verbindung abgedeckt ist, die Polymerisation am 5'- oder 3'-Ende verhindert, oder Kombinationen derselben ein. Die Komplementarität eines Nukleinsäurefragments zu einem ausgewählten Target-Nukleinsäuremolekül bedeutet im allgemeinen, daß wenigstens etwa 70% spezifische Basenpaarung über die gesamte Länge des Fragments gezeigt wird. Vorzugsweise zeigt das Nukleinsäurefragment wenigstens etwa 80% spezifische Basenpaarung; und am bevorzugtesten wenigstens etwa 90%. Assays zum Bestimmen des prozentualen Mismatch (und somit der prozentualen spezifischen Basenpaarung) sind im Stand der Technik gut bekannt und beruhen auf dem prozentualen Mismatch als einer Funktion des T<sub>m</sub>, wenn bezogen auf eine Kontrolle mit vollständiger Basenpaarung.

[0072] Wie hierin verwendet, bezieht sich der Begriff "Alkyl", allein oder in Kombination, auf einen gesättigten, geradekettigen oder verzweigt-kettigen Kohlenwasserstoffrest, der von 1 bis 10, vorzugsweise von 1 bis 6 und bevorzugter von 1 bis 4 Kohlenstoffatome enthält. Beispiele für solche Reste schließen Methyl, Ethyl, n-Propyl, iso-Propyl, n-Butyl, iso-Butyl, sec.-Butyl, tert.-Butyl, Pentyl, iso-Amyl, Hexyl, Decyl und dergleichen ein, sind aber nicht hierauf beschränkt. Der Begriff "Alkylen" bezieht sich auf einen gesättigten, geradkettigen oder verzweigt-kettigen Kohlenwasserstoffdirekt, der von 1 bis 10, vorzugsweise von 1 bis 6 und bevorzugter von 1 bis 4 Kohlenstoffe enthält. Beispiele für solche einen Direkt schließen Methylen, Ethylen (–CH<sub>2</sub>–CH<sub>2</sub>–), Propylen und dergleichen ein, sind aber nicht hierauf beschränkt.

[0073] Der Begriff "Alkenyl", allein oder in Kombination, bezieht sich auf einen geradkettigen oder verzweigt-kettigen Kohlenwasserstoffrest mit wenigstens einer Kohlenstoff-Kohlenstoff-Doppelbindung in insgesamt von 2 bis 10, vorzugsweise von 2 bis 6 und bevorzugter von 2 bis 4 Kohlenstoffatomen. Beispiele für solche Reste schließen Ethenyl, E- und Z-Propenyl, Isopropenyl, E- und Z-Butenyl, E- und Z-Isobutenyl, E- und Z-Pentenyl, Decenyl und dergleichen ein, sind aber nicht hierauf beschränkt. Der Begriff "Alkenylen" bezieht sich auf einen geradkettigen oder verzweigt-kettigen Kohlenwasserstoffdirekt mit wenigstens einer Kohlenstoff-Kohlenstoff-Doppelbindung in insgesamt von 2 bis 10, vorzugsweise von 2 bis 6 und bevorzugter von 2 bis 4 Kohlenstoffatomen. Beispiele für solche Direkte schließen Methylen (=CH<sub>2</sub>), Ethylen (–CH=CH–), Propylen (–CH<sub>2</sub>–CH=CH–) und dergleichen ein, sind aber nicht hierauf beschränkt.

[0074] Der Begriff "Alkynyl", allein oder in Kombination, bezieht sich auf einen geradkettigen oder verzweigt-kettigen Kohlenwasserstoffrest mit wenigstens einer Kohlenstoff-Kohlenstoff-Dreifachbindung in insgesamt



von 2 bis 10, vorzugsweise von 2 bis 6 und bevorzugter von 2 bis 4 Kohlenstoffatomen. Beispiele für solche Reste schließen Ethinyl (Acetylenyl), Propinyl (Propargyl), Butinyl, Hexinyl, Decinyl und dergleichen ein, sind aber nicht hierauf beschränkt. Der Begriff "Alkinylen", allein oder in Kombination, bezieht sich auf einen geradkettigen oder verzweigt-kettigen Kohlenwasserstoffdirekt mit wenigstens einer Kohlenstoff-Kohlenstoff-Dreifachbindung in insgesamt von 2 bis 10, vorzugsweise von 2 bis 6 und bevorzugter von 2 bis 4 Kohlenstoffatomen. Beispiele für solche Reste schließen Ethinylen ( $-C\equiv C-$ ), Propinylen ( $-CH_2-C\equiv C-$ ) und dergleichen ein, sind aber nicht hierauf beschränkt.

[0075] Der Begriff "Cycloalkyl", allein oder in Kombination, bezieht sich auf eine gesättigte cyclische Anordnung von Kohlenstoffatomen, mit einer Anzahl von 3 bis 8 und vorzugsweise von 3 bis 6 Kohlenstoffatomen. Beispiele für solche Cycloalkylreste schließen Cyclopropyl, Cyclobutyl, Cyclopentyl, Cyclohexyl und dergleichen ein, sind aber nicht hierauf beschränkt. Der Begriff "Cycloalkylen" bezeichnet eine Direktform eines Cycloalkyls.

[0076] Der Begriff "Cycloalkenyl", allein oder in Kombination, bezieht sich auf einen cyclischen Kohlenstoffring, der von 4 bis 8, vorzugsweise 5 oder 6, Kohlenstoffatome und eine oder mehrere Doppelbindungen enthält. Beispiele für solche Cycloalkenylreste schließen Cyclopentenyl, Cyclohexenyl, Cyclopentadienyl und dergleichen ein, sind aber nicht hierauf beschränkt. Der Begriff "Cycloalkenylen" bezieht sich auf eine Direktform eines Cycloalkenyls.

[0077] Der Begriff "Aryl" bezieht sich auf eine carbocyclische (vollständig aus Kohlenstoff und Wasserstoff bestehende) aromatische Gruppe, die ausgewählt ist aus der Gruppe, die aus Phenyl, Naphthyl, Indenyl, Indanyl, Azulenyl, Fluorenyl und Anthracenyl besteht; oder eine heterocyclische aromatische Gruppe, die ausgewählt ist aus der Gruppe, die aus Furyl, Thienyl, Pyridyl, Pyrrolyl, Oxazolyl, Thiazolyl, Imidazolyl, Pyrazolyl, 2-Pyrazolyl, Pyrazolidinyl, Isoxazolyl, Isothiazolyl, 1,2,3-Oxadiazolyl, 1,2,3-Triazolyl, 1,3,4-Thiadiazolyl, Pyridazinyl, Pyrimidinyl, Pyrazinyl, 1,3,5-Triazinyl, 1,3,5-Trithianyl, Indolizinyll, Indolyl, Isoindolyl, 3H-Indolyl, Indolinyl, Benzo[b]furanlyl, 2,3-Dihydrobenzofuranlyl, Benzo[b]thiophenyl, 1H-Indazolyl, Benzimidazolyl, Benzthiazolyl, Purinyl, 4H-Chinolinyl, Chinolinyl, Isochinolinyl, Cinnolinyl, Phthalazinyl, Chinazolinyl, Chinoxalinyll, 1,8-Naphthyridinyl, Pteridinyl, Carbazolyl, Acridinyl, Phenazinyl, Phenothiazinyl und Phenoxazinyl besteht.

[0078] "Aryl"-Gruppen, wie in dieser Anmeldung definiert, können unabhängig einen bis vier Substituenten enthalten, die unabhängig ausgewählt sind aus der Gruppe, die aus Wasserstoff, Halogen, Hydroxyl, Amino, Nitro, Trifluormethyl, Trifluormethoxy, Alkyl, Alkenyl, Alkinyll, Cyano, Carboxy, Carboalkoxy, 1,2-Dioxyethylen, Alkoxy, Alkenoxy oder Alkinoxy, Alkylamino, Alkenylamino, Alkinyllamino, aliphatischem oder aromatischem Acyl, Alkoxy-carbonylamino, Alkylsulfonylamino, Morpholinocarbonylamino, Thiomorpholinocarbonylamino, N-Alkylguanidino, Aralkylaminosulfonyl; Aralkoxyalkyl; N-Aralkoxyharnstoff; N-Hydroxylharnstoff; N-Alkenylharnstoff; N,N-(Alkyl, Hydroxyl)harnstoff Heterocyclyl; Thioaryloxy-substituiertem Aryl; N,N-(Aryl, Alkyl)hydrazino; Ar'-substituiertem Sulfonylheterocyclyl; Aralkyl-substituiertem Heterocyclyl; Cycloalkyl- und Cycloalkenyl-substituiertem Heterocyclyl; Cycloalkyl-kondensiertem Aryl; Aryloxy-substituiertem Alkyl; Heterocyclylamino; aliphatischem oder aromatischem Acylaminocarbonyl; aliphatischem oder aromatischem Acyl-substituiertem Alkenyl; Ar'-substituiertem Aminocarbonyloxy; Ar', Ar'-disubstituiertem Aryl; aliphatischem oder aromatischem Acyl-substituiertem Acyl; Cycloalkylcarbonylalkyl; Cycloalkyl-substituiertem Amino; Aryloxy-carbonylalkyl; Phosphorodiamidylsäure oder -ester besteht.

[0079] "Ar" ist eine carbocyclische oder heterocyclische Arylgruppe, wie oben definiert, mit einem bis drei Substituenten, die ausgewählt sind aus der Gruppe, die aus Wasserstoff, Halogen, Hydroxyl, Amino, Nitro, Trifluormethyl, Trifluormethoxy, Alkyl, Alkenyl, Alkinyll, 1,2-Dioxy-methylen, 1,2-Dioxyethylen, Alkoxy, Alkenoxy, Alkinoxy, Alkylamino, Alkenylamino oder Alkinyllamino, Alkylcarbonyloxy, aliphatischem oder aromatischem Acyl, Alkylcarbonylamino, Alkoxy-carbonylamino, Alkylsulfonylamino, N-Alkyl- oder N,N-Dialkylharnstoff besteht.

[0080] Der Begriff "Alkoxy", allein oder in Kombination, bezieht sich auf ein Alkyletherrest, wobei der Begriff "Alkyl" wie oben definiert ist. Beispiele für geeignete Alkyletherreste schließen Methoxy, Ethoxy, n-Propoxy, iso-Propoxy, n-Butoxy, iso-Butoxy, sec.-Butoxy, tert.-Butoxy und dergleichen ein, sind aber nicht hierauf beschränkt.

[0081] Der Begriff "Alkenoxy", allein oder in Kombination, bezieht sich auf einen Rest der Formel Alkenyl-O-, wobei der Begriff "Alkenyl" wie oben definiert ist, vorausgesetzt, daß der Rest kein Enoether ist. Beispiele für geeignete Alkenoxyreste schließen Allyloxy, E- und Z-3-Methyl-2-propenoxy und dergleichen ein, sind aber nicht hierauf beschränkt.

[0082] Der Begriff "Alkinyloxy", allein oder in Kombination, bezieht sich auf einen Rest der Formel Alkinyll-O-, wobei der Begriff "Alkinyll" wie oben definiert ist, vorausgesetzt, daß der Rest kein Inoether ist. Beispiele für geeignete Alkinyloxyreste schließen Propargyloxy, 2-Butinyloxy und dergleichen ein, sind aber nicht hierauf beschränkt.

[0083] Der Begriff "Thioalkoxy" bezieht sich auf einen Thioetherrest der Formel Alkyl-S-, wobei Alkyl wie oben definiert ist.

[0084] Der Begriff "Alkylamino", allein oder in Kombination, bezieht sich auf einen Mono- oder Dialkyl-substi-

tuerten Aminorest (d. h. einen Rest der Formel Alkyl-NH- oder (Alkyl)<sub>2</sub>-N-), wobei der Begriff "Alkyl" wie oben definiert ist. Beispiele für geeignete Alkylaminoreste schließen Methylamino, Ethylamino, Propylamino, Isopropylamino, tert.-Butylamino, N,N-Diethylamino und dergleichen ein, sind aber nicht hierauf beschränkt.

[0085] Der Begriff "Alkenylamino", allein oder in Kombination, bezieht sich auf einen Rest der Formel Alkenyl-NH- oder (Alkenyl)<sub>2</sub>-N-, wobei der Begriff "Alkenyl" wie oben definiert ist, vorausgesetzt, daß der Rest kein Enamin ist. Ein Beispiel für solche Alkenylaminoreste ist der Allylaminorest.

[0086] Der Begriff "Alkynylamino", allein oder in Kombination, bezieht sich auf einen Rest der Formel Alkynyl-NH- oder (Alkynyl)<sub>2</sub>-N-, in der der Begriff "Alkynyl" wie oben definiert ist, vorausgesetzt, daß der Rest kein Inamin ist. Ein Beispiel für solche Alkynylaminoreste ist der Propargylaminorest.

[0087] Der Begriff "Amid" bezieht sich entweder auf -N(R<sup>1</sup>)-C(=O)- oder -C(=O)-N(R<sup>1</sup>)-, wobei R<sup>1</sup> hierin so definiert ist, daß es Wasserstoff sowie andere Gruppen einschließt. Der Begriff "substituiertes Amid" bezieht sich auf die Situation, in der R<sup>1</sup> nicht Wasserstoff ist, während der Begriff "unsubstituiertes Amid" sich auf die Situation bezieht, in der R<sup>1</sup> Wasserstoff ist.

[0088] Der Begriff "Aryloxy", allein oder in Kombination, bezieht sich auf einen Rest der Formel Aryl-O-, wobei Aryl wie oben definiert ist. Beispiele für Aryloxyreste schließen Phenoxy, Naphthoxy, Pyridyloxy und dergleichen ein, sind aber nicht hierauf beschränkt.

[0089] Der Begriff "Arylamino", allein oder in Kombination, bezieht sich auf einen Rest der Formel Aryl-NH-, in der Aryl wie oben definiert ist. Beispiele für Arylaminoreste schließen Phenylamino (Anilido), Naphthylamino, 2-, 3- und 4-Pyridylamino und dergleichen ein, sind aber nicht hierauf beschränkt.

[0090] Der Begriff "Aryl-kondensiertes Cycloalkyl", allein oder in Kombination, bezieht sich auf einen Cycloalkylrest, der zwei benachbarte Atome mit einem Arylrest teilt, wobei die Begriffe "Cycloalkyl" und "Aryl" wie oben definiert sind. Ein Beispiel eines Aryl-kondensierten Cycloalkylrestes ist der Benzo-kondensierte Cyclobutylrest.

[0091] Der Begriff "Alkylcarbonylamino", allein oder in Kombination, bezieht sich auf einen Rest der Formel Alkyl-CONH-, wobei der Begriff "Alkyl" wie oben definiert ist.

[0092] Der Begriff "Alkoxy-carbonylamino", allein oder in Kombination, bezieht sich auf einen Rest der Formel Alkyl-OC(=O)NH-, wobei der Begriff "Alkyl" wie oben definiert ist.

[0093] Der Begriff "Alkylsulfonylamino", allein oder in Kombination, bezieht sich auf einen Rest der Formel Alkyl-SO<sub>2</sub>NH-, wobei der Begriff "Alkyl" wie oben definiert ist.

[0094] Der Begriff "Arylsulfonylamino", allein oder in Kombination, bezieht sich auf einen Rest der Formel Aryl-SO<sub>2</sub>NH-, wobei der Begriff "Aryl" wie oben definiert ist.

[0095] Der Begriff "N-Alkylharnstoff", allein oder in Kombination, bezieht sich auf einen Rest der Formel Alkyl-NH-CO-NH-, wobei der Begriff "Alkyl" wie oben definiert ist.

[0096] Der Begriff "N-Arylharnstoff", allein oder in Kombination, bezieht sich auf einen Rest der Formel Aryl-NH-CO-NH-, wobei der Begriff "Aryl" wie oben definiert ist.

[0097] Der Begriff "Halogen" bedeutet Fluor, Chlor, Brom und Iod.

[0098] Der Begriff "Kohlenwasserstoffrest" bezieht sich auf eine Anordnung von Kohlenstoff- und Wasserstoffatomen, die nur ein einziges Wasserstoffatom benötigt, um ein unabhängiges stabiles Molekül zu sein. Somit hat ein Kohlenwasserstoffrest eine offene Valenzstelle an einem Kohlenstoffatom, durch die der Kohlenwasserstoffrest an ein anderes Atom (andere Atome) gebunden werden kann. Alkyl, Alkenyl, Cycloalkyl, etc. sind Beispiele für Kohlenwasserstoffreste.

[0099] Der Begriff "Kohlenwasserstoffdirekt" bezieht sich auf eine Anordnung von Kohlenstoff- und Wasserstoffatomen, die zwei Wasserstoffatome benötigt, um ein unabhängiges stabiles Molekül zu sein. Somit hat ein Kohlenwasserstoffrest zwei offene Valenzstellen an einem oder zwei Kohlenstoffatomen, durch die der Kohlenwasserstoffrest an ein anderes Atom (andere Atome) gebunden werden kann. Alkylen, Alkenylen, Alkinylen, Cycloalkylen, etc. sind Beispiele für Kohlenwasserstoffdirekte.

[0100] Der Begriff "Hydrocarbyl" bezieht sich auf jede stabile Anordnung, die vollständig aus Kohlenstoff und Wasserstoff besteht, mit einer einzelnen Valenzstelle, an der sie an eine andere Einheit gebunden ist, und schließt somit Reste ein, die als Alkyl, Alkenyl, Alkynyl, Cycloalkyl, Cycloalkenyl, Aryl (ohne Einbeziehung von Heteroatomen in den Arylring), Arylalkyl, Alkylaryl und dergleichen ein. Kohlenwasserstoffrest ist ein anderer Name für Hydrocarbyl.

[0101] Der Begriff "Hydrocarbylen" bezieht sich auf jede stabile Anordnung, die vollständig aus Kohlenstoff und Wasserstoff besteht, mit zwei Valenzstellen, an der sie an andere Einheiten gebunden ist, und schließt somit Alkylen, Alkenylen, Alkinylen, Cycloalkylen, Cycloalkenylen, Arylen (ohne Einbeziehung von Heteroatomen in den Arylring), Arylalkylen, Alkylarylen und dergleichen ein. Kohlenwasserstoffdirekt ist ein anderer Name für Hydrocarbylen.

[0102] Der Begriff "Hydrocarbyl-O-hydrocarbylen" bezieht sich auf eine Hydrocarbyl-Gruppe, die an ein Sauerstoffatom gebunden ist, wobei das Sauerstoffatom in ähnlicher Weise an eine Hydrocarbylen-Gruppe an einer der zwei Valenzstellen gebunden ist, an der die Hydrocarbylen-Gruppe an andere Einheiten gebunden ist. Die Begriffe "Hydrocarbyl-S-hydrocarbylen", "Hydrocarbyl-NH-hydrocarbylen" und "Hydrocarbyl-Amid-hydro-

carbylen" haben äquivalente Bedeutungen, wobei Sauerstoff durch Schwefel, -NH- bzw. eine Amidgruppe ersetzt worden ist.

[0103] Der Begriff N-(Hydrocarbyl)hydrocarbylen bezieht sich auf eine Hydrocarbylen-Gruppe, wobei eine der zwei Valenzstellen an ein Stickstoffatom gebunden ist und dieses Stickstoffatom gleichzeitig an einen Wasserstoff und eine Hydrocarbyl-Gruppe gebunden ist.

[0104] Der Begriff N,N-Di(hydrocarbyl)hydrocarbylen bezieht sich auf eine Hydrocarbylengruppe, wobei eine der zwei Valenzstellen an ein Stickstoffatom gebunden ist und dieses Stickstoffatom gleichzeitig an zwei Hydrocarbyl-Gruppen gebunden ist.

[0105] Der Begriff "Hydrocarbylacyl-hydrocarbylen" bezieht sich auf eine Hydrocarbyl-Gruppe, die durch eine Acyl(-C(=O)-)-Gruppe an eine der zwei Valenzstellen einer Hydrocarbylen-Gruppe gebunden ist.

[0106] Die Begriffe "Heterocyclhydrocarbyl" und "Heterocycl" beziehen sich auf eine stabile cyclische Anordnung von Atomen, die Kohlenstoffatome und bis zu vier Atome (als Heteroatome bezeichnet) einschließt, die ausgewählt sind aus Sauerstoff, Stickstoff, Phosphor und Schwefel. Die cyclische Anordnung kann in der Form eines monocyclischen Ringes mit 3–7 Atomen oder eines bicyclischen Ringes mit 8–11 Atomen vorliegen. Die Ringe können gesättigt oder ungesättigt sein (einschließlich aromatischer Ringe) und können fakultativ Benzo-kondensiert sein. Stickstoff- und Schwefelatomen im Ring können in jeder oxidierten Form vorliegen, einschließlich der quaternisierten Form von Stickstoff. Ein Heterocyclhydrocarbyl kann an irgendeinem endocyclischen Kohlenstoff oder Heteroatom gebunden sein, das zur Schaffung einer stabilen Struktur führt. Bevorzugte Heterocyclhydrocarbyle schließen 5- bis 7-gliedrige monocyclische Heterocyclen ein, die ein oder zwei Stickstoff-Heteroatome enthalten.

[0107] Ein substituiertes Heterocyclhydrocarbyl bezieht sich auf ein Heterocyclhydrocarbyl, wie oben definiert, wobei wenigstens ein Ringatom davon an einen angegebenen Substituenten gebunden ist, der sich vom Ring weg erstreckt.

[0108] Bei Bezugnahme auf Hydrocarbyl- und Hydrocarbylen-Gruppen bezieht sich der Begriff "Derivate von irgendeinem der vorstehenden, wobei ein oder mehr Wasserstoffe durch eine gleiche Anzahl Fluoride ersetzt ist" auf Moleküle, die Kohlenstoff-, Wasserstoff- und Fluoridatome enthalten, aber keine anderen Atome.

[0109] Der Begriff "aktivierter Ester" ist ein Ester, der eine "Abgangsgruppe" enthält, die leicht durch ein Nukleophil verdrängbar ist, wie etwa ein Amin und Alkohol oder ein Thiol-Nukleophil. Solche Abgangsgruppen sind gut bekannt und schließen, ohne Beschränkung, N-Hydroxysuccinimid, N-Hydroxybenzotriazol, Halogen (Halogenide), Alkoxy einschließlich Tetrafluorphenolaten, Thioalkoxy und dergleichen ein. Der Begriff "geschützter Ester" bezieht sich auf eine Estergruppe, die maskiert oder in anderer Weise unreaktiv ist. Siehe z. B. Greene, "Protecting Groups In Organic Synthesis".

[0110] Angesichts der obigen Definitionen können weitere chemische Begriffe, die in dieser Anmeldung verwendet werden, von den Fachleuten leicht verstanden werden. Begriffe können allein oder in irgendeiner Kombination derselben verwendet werden. Die bevorzugten und weiter bevorzugten Kettenlängen der Reste gelten für alle solche Kombinationen.

## A. ERZEUGUNG MARKIERTER NUKLEINSÄUREFRAGMENTE

[0111] Wie oben angegeben, stellt ein Aspekt der vorliegenden Erfindung ein allgemeines Schema zur DNA-Sequenzierung bereit, das die Verwendung von mehr als 16 Markierungen in jeder Spur ermöglicht; bei kontinuierlichem Nachweis können die Markierungen nachgewiesen und die Sequenz gelesen werden, wenn die Trennung nach Größe abläuft, ebenso wie bei herkömmlicher Sequenzierung auf Fluoreszenz-Basis. Dieses Schema ist anwendbar auf alle DNA-Sequenzierungstechniken, die auf Trennung der markierten Moleküle nach Größe beruhen. Geeignete Markierungen und Linker zur Verwendung in der vorliegenden Erfindung sowie Verfahren zur Sequenzierung von Nukleinsäuren sind unten detaillierter diskutiert.

### 1. Markierungen

[0112] "Markierung", wie hierin verwendet, bezieht sich allgemein auf eine chemische Einheit, die verwendet wird, um ein "interessierendes Molekül" in eindeutiger Weise zu identifizieren, und bezieht sich, genauer gesagt, auf die variable Markierungskomponente sowie auf alles, was in dem Markierungs-Reaktanten, der Markierungs-Komponente und/oder der Markierungs-Einheit am dichtesten daran gebunden ist.

[0113] Eine Markierung, die in der vorliegenden Erfindung nützlich ist, besitzt mehrere Eigenschaften:

- 1) Sie kann von allen anderen Markierungen unterschieden werden. Diese Unterscheidung von anderen chemischen Einheiten basiert auf ihren spektroskopischen Eigenschaften d. h. in der Massenspektroskopie.
- 2) Die Markierung kann nachgewiesen werden, wenn sie mit  $10^{-22}$  bis  $10^{-6}$  Mol vorliegt.
- 3) Die Markierung besitzt einen chemischen Griff, durch den sie an das MOI gebunden werden kann, welches die Markierung in eindeutiger Weise identifizieren soll. Die Bindung kann direkt an das MOI erfolgen oder indirekt durch eine "Linker"-Gruppe.

4) Die Markierung ist chemisch stabil gegenüber allen Manipulationen, denen sie unterworfen wird, einschließlich Bindung und Spaltung vom MOI, und allen Manipulationen des MOI, während die Markierung daran gebunden ist.

5) Die Markierung beeinflusst die Manipulationen, die am MOI durchgeführt werden, während die Markierung daran gebunden ist, nicht in signifikanter Weise. Wenn die Markierung z. B. an ein Oligonukleotid gebunden ist, darf die Markierung irgendwelche Hybridisierungs- oder enzymatischen Reaktionen (z. B. PCR-Sequenzierungsreaktionen), die am Nukleotid durchgeführt werden, nicht signifikant beeinflussen. In ähnlicher Weise darf, wenn die Markierung an einen Antikörper gebunden ist, sie die Antigen-Erkennung durch den Antikörper nicht signifikant beeinflussen.

[0114] Eine Markierungs-Einheit, die mit einem bestimmten spektroskopischen oder potentiometrischen Verfahren nachgewiesen werden soll, sollte Eigenschaften besitzen, die die Empfindlichkeit und Spezifität des Nachweises mit diesem Verfahren erhöht. Typischerweise wird die Markierungs-Einheit diese Eigenschaften haben, weil sie in die variable Markierungs-Komponente hineinkonstruiert worden sind, die typischerweise den Hauptteil der Markierungs-Einheit darstellen wird. In der folgenden Diskussion bezieht sich die Verwendung des Wortes "Markierung" typischerweise auf die Markierungs-Einheit (d. h. das Spaltprodukt, das die variable Markierungs-Komponente enthält), kann jedoch auch so betrachtet werden, daß es die variable Markierungs-Komponente selbst bezeichnet, weil diese der Teil der Markierungs-Einheit ist, der typischerweise dafür verantwortlich ist, die eindeutig nachweisbaren Eigenschaften bereitzustellen. In Verbindungen der Formel T-L-X wird der "T"-Teil die variable Markierungs-Komponente enthalten, der mit Massenspektrometrie charakterisiert wird, der "T"-Teil von T-L-X wird daher als  $T^{ms}$  bezeichnet. In ähnlicher Weise kann das Spaltprodukt von T-L-X, das T enthält, als die  $T^{ms}$ -enthaltende Einheit bezeichnet werden. Die folgenden spektroskopischen und potentiometrischen Verfahren können verwendet werden, um  $T^{ms}$ -enthaltende Einheiten zu charakterisieren.

#### a. Eigenschaften von MS-Markierungen

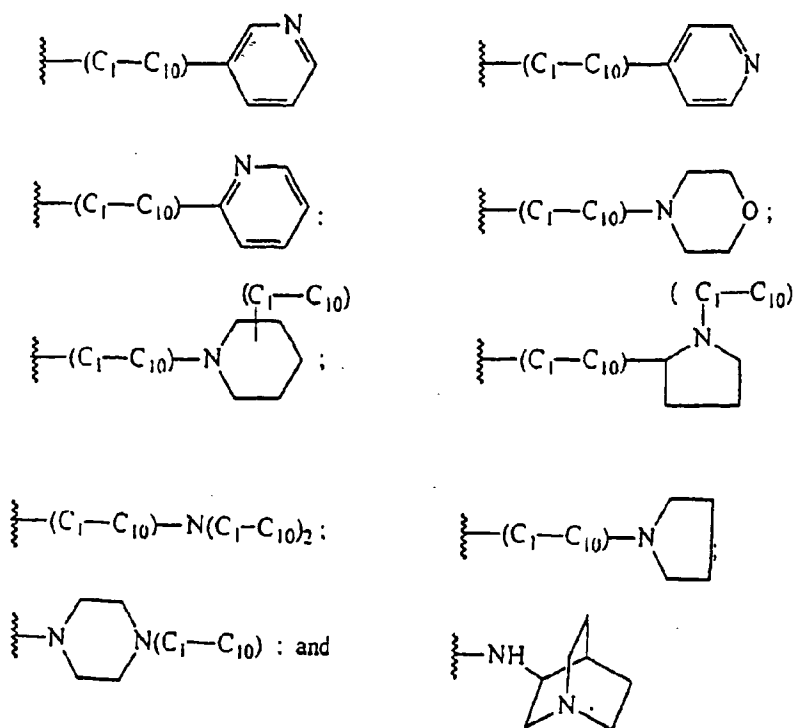
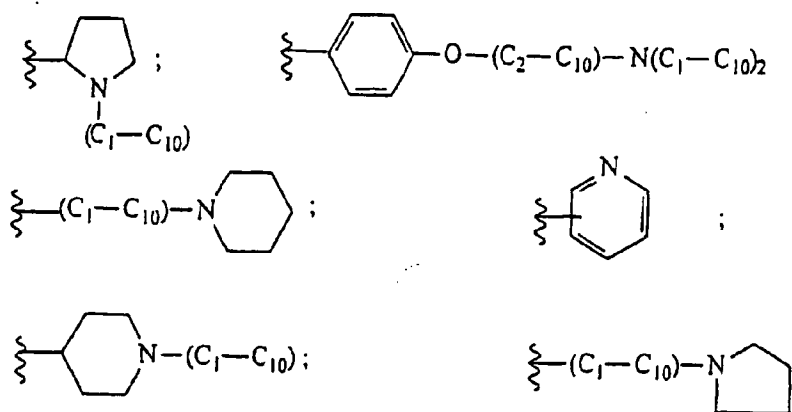
[0115] Der Marker ist durch Massenspektrometrie analysierbar (d. h. ist eine MS-lesbare Markierung, hierin auch als eine MS-Markierung oder " $T^{ms}$ -enthaltende Einheit" bezeichnet). Das wesentliche Merkmal der Markierung ist, daß sie ionisiert werden kann. Es ist somit ein bevorzugtes Element bei der Konstruktion von MS-lesbaren Markierungen, darin eine chemische Funktionalität einzubauen, die unter Ionisationsbedingungen in MS eine positive oder negative Ladung tragen kann. Dieses Merkmal verleiht verbesserte Effizienz der Ionenbildung und höhere Gesamttempfindlichkeit des Nachweises, insbesondere bei Elektrosprayionisation. Die chemische Funktionalität, die eine ionisierte Ladung unterstützt, kann sich von  $T^{ms}$  oder L oder beiden ableiten. Faktoren, die die relative Empfindlichkeit eines durch Massenspektrometrie nachzuweisenden Analyten erhöhen kann, sind in z. B. Sunner, J., et al., Anal. Chem. 60: 1300-1307 (1988), diskutiert.

[0116] Eine bevorzugte Funktionalität, um das Tragen einer negativen Ladung zu erleichtern, ist eine organische Säure, wie etwa phenolisches Hydroxyl, Carbonsäure, Phosphonat, Phosphat, Tetrazol, Sulfonylharnstoff Perfluoralkohol und Sulfonsäure.

[0117] Bevorzugte Funktionalität, um das Tragen einer positiven Ladung unter Ionisationsbedingungen zu erleichtern, sind aliphatische oder aromatische Amine. Beispiele für funktionelle Amingruppen, die erhöhte Nachweisbarkeit von MS-Markierungen ergeben, schließen quartäre Amine (d. h. Amine, die vier Bindungen, jede zu Kohlenstoffatomen, aufweisen, siehe Aebersold, U.S.-Patent Nr. 5,240,859) und tertiäre Amine ein (d. h. Amine, die drei Bindungen, jede zu Kohlenstoffatomen, aufweisen, was C=N-C-Gruppen einschließt, wie sie in Pyridin vorhanden sind, siehe Hess et al., Anal. Biochem. 224: 373, 1995; Bures et al., Anal. Biochem. 224-364, 1995). Gehinderte tertiäre Amine sind besonders bevorzugt. Tertiäre und quartäre Amine können Alkyl oder Aryl sein. Eine  $T^{ms}$ -enthaltende Einheit muß wenigstens eine ionisierbare Spezies tragen, kann aber mehr als eine ionisierbare Spezies besitzen. Der bevorzugte Ladungszustand ist eine einzelne ionisierte Spezies pro Markierung.

[0118] Demgemäß ist es bevorzugt, daß jede  $T^{ms}$ -enthaltende Einheit (und jede variable Markierungs-Komponente) nur ein einziges gehindertes Amin oder eine einzige organische Säuregruppe enthält.

[0119] Geeignete Amin-enthaltende Reste, die Teil der  $T^{ms}$ -enthaltenden Einheit bilden können, schließen die folgenden ein:



[0120] Die Identifizierung einer Markierung durch Massenspektrometrie beruht vorzugsweise auf ihrem Verhältnis Molekülmasse zu Ladung ( $m/z$ ). Der bevorzugte Molekülmassenbereich für MS-Markierungen ist von etwa 100 bis 2.000 Daltons, und vorzugsweise hat die  $T^{\text{ms}}$ -enthaltende Einheit eine Masse von wenigstens etwa 250 Daltons, bevorzugter wenigstens etwa 300 Daltons und noch bevorzugter wenigstens etwa 350 Daltons. Es ist im allgemeinen schwierig für Massenspektrometer, zwischen Einheiten zu unterscheiden, die Moleküle unter etwa 200–250 Daltons haben (in Abhängigkeit von dem genauen Instrument), und so haben bevorzugte  $T^{\text{ms}}$ -enthaltende Einheiten der Erfindung Massen oberhalb dieses Bereiches.

[0121] Wie oben erklärt, kann die  $T^{\text{ms}}$ -enthaltende Einheit andere Atome als diejenigen enthalten, die in der variablen Markierungs-Komponente vorhanden sind und tatsächlich andere als in  $T^{\text{ms}}$  selbst vorhanden sind. Demgemäß kann die Masse von  $T^{\text{ms}}$  selbst niedriger sein als etwa 250 Daltons, so lange die  $T^{\text{ms}}$ -enthaltende Einheit eine Masse von wenigstens etwa 250 Daltons hat. So kann die Masse von  $T^{\text{ms}}$  von 15 (d. h. einem Methylrest) bis etwa 10.000 Daltons reichen und reicht vorzugsweise von etwa 100 bis etwa 5.000 Daltons und reicht bevorzugter von etwa 200 bis etwa 1.000 Daltons.

[0122] Es ist relativ schwierig, Markierungen durch Massenspektrometrie zu unterscheiden, wenn diese Mar-

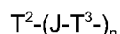
kierungen Atome enthalten, die mehr als ein Isotop in signifikanter Menge aufweisen. Demgemäß enthalten T-Gruppen, die für massenspektroskopische Identifizierung ( $T^{ms}$ -Gruppen) gedacht sind, Kohlenstoff, wenigstens eines von Wasserstoff und Fluorid und fakultative Atome, die ausgewählt sind aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Iod. Obgleich weitere Atome im  $T^{ms}$  vorhanden sein können, kann ihr Vorhandensein die Analyse der Massenspektrometerdaten etwas schwieriger machen. Vorzugsweise haben die  $T^{ms}$ -Gruppen nur Kohlenstoff-, Stickstoff- und Sauerstoffatome, zusätzlich zu Wasserstoff und/oder Fluorid.

[0123] Fluorid ist ein fakultatives, aber bevorzugtes Atom für die Einbeziehung in eine  $T^{ms}$ -Gruppe. Im Vergleich zu Wasserstoff ist Fluorid natürlich viel schwerer. So führt das Vorhandensein von Fluoridatomen statt Wasserstoffatomen zu  $T^{ms}$ -Gruppen mit höherer Masse, wodurch ermöglicht wird, daß die  $T^{ms}$ -Gruppe eine Masse von mehr als 250 Daltons erreicht und übersteigt, was wünschenswert ist, wie oben erläutert. Zusätzlich verleiht der Ersatz von Wasserstoff durch Fluorid der  $T^{ms}$ -enthaltenden Einheit größere Flüchtigkeit, und größere Flüchtigkeit des Analyten erhöht die Empfindlichkeit, wenn Massenspektrometrie als das Nachweisverfahren verwendet wird.

[0124] Die Molekülformel von  $T^{ms}$  fällt in den Bereich  $C_{1-500}N_{0-100}O_{0-100}P_{0-10}H_{\alpha}F_{\beta}I_{\delta}$ , wobei die Summe von  $\alpha$ ,  $\beta$  und  $\delta$  ausreichend ist, um die sonst ungesättigten Valenzen der C-, N-, O-, S- und P-Atome abzusättigen. Die Bezeichnung  $C_{1-500}N_{0-100}O_{0-100}P_{0-10}H_{\alpha}F_{\beta}I_{\delta}$  bedeutet, daß  $T^{ms}$  wenigstens ein Kohlenstoffatom enthält und jede Zahl von 1 bis 500 Kohlenstoffatome enthalten kann, zusätzlich dazu, daß es so viel wie 100 Stickstoffatome (" $N_0$ ." bedeutet, daß  $T^{ms}$  keinerlei Stickstoffatome enthalten muß), und so viel wie 100 Sauerstoffatome und so viel wie 10 Schwefelatome und so viel wie 10 Phosphoratom enthält. Die Symbole  $\alpha$ ,  $\beta$  und  $\delta$  stellen die Anzahl von Wasserstoff-, Fluorid- und Iodidatomen in  $T^{ms}$  dar, wobei jeweils zwei dieser Zahlen Null sein können und wobei die Summe dieser Zahlen der Gesamtheit der sonst ungesättigten Valenzen der C-, N-, O-, S- und P-Atome entspricht. Vorzugsweise hat  $T^{ms}$  eine Molekülformel, die in den Bereich  $C_{1-50}N_{0-10}O_{0-10}H_{\alpha}F_{\beta}$  fällt, wobei die Summe von  $\alpha$  und  $\beta$  der Anzahl von Wasserstoff- bzw. Fluoridatomen entspricht, die in der Einheit vorhanden sind.

[0125] Zusätzlich zu den obigen Eigenschaften ist es bevorzugt, daß die Markierung eine modulare chemische Struktur besitzt. Dies hilft bei der Konstruktion einer großen Anzahl von strukturell verwandten Markierungen unter Verwendung der Techniken der kombinatorischen Chemie. Die  $T^{ms}$ -Gruppe hat z. B. wünschenswerterweise mehrere Eigenschaften. Sie enthält wünschenswerterweise eine funktionelle Gruppe, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die  $T^{ms}$ -enthaltende Einheit Massenspektrometrie unterzogen wird (einfacher als eine "Massenspektrometrieempfindlichkeitsverstärker"-Gruppe oder MSSE bezeichnet). Auch kann sie wünschenswerterweise als ein Mitglied in einer Familie von  $T^{ms}$ -enthaltenden Einheiten dienen, in der die Mitglieder der Familie jedes ein unterschiedliches Masse/Ladungs-Verhältnis haben, jedoch ungefähr dieselbe Empfindlichkeit im Massenspektrometer aufweisen. Mitglieder der Familie haben somit wünschenswerterweise denselben MSSE. Um die Schaffung von Familien von Verbindungen zu ermöglichen, hat es sich als geeignet erwiesen, Markierungs-Reaktanten über ein modulares Syntheschema zu erzeugen, so daß die Markierungs-Komponenten selbst als Module-umfassend angesehen werden können.

[0126] In einem bevorzugten modularen Ansatz für die Struktur der  $T^{ms}$ -Gruppe hat  $T^{ms}$  die Formel



wobei  $T^2$  eine organische Einheit ist, die aus Kohlenstoff und einem oder mehreren von Wasserstoff, Fluorid, Iodid, Sauerstoff, Stickstoff, Schwefel und Phosphor gebildet wird, mit einem Massebereich von 15 bis 500 Daltons;  $T^3$  eine organische Einheit ist, die aus Kohlenstoff und einem oder mehreren von Wasserstoff, Fluorid, Iodid, Sauerstoff, Stickstoff, Schwefel und Phosphor gebildet wird, mit einem Massebereich von 50 bis 1.000 Daltons; J eine direkte Bindung oder eine funktionelle Gruppe ist, wie etwa Amid, Ester, Amin, Sulfid, Ether, Thioester, Disulfid, Thioether, Harnstoff, Thioharnstoff, Carbamat, Thiocarbamat, Schiff'sche Base, reduzierte Schiff'sche Base, Imin, Oxim, Hydrazon, Phosphat, Phosphonat, Phosphoramid, Phosphonamid, Sulfonat, Sulfonamid oder Kohlenstoff-Kohlenstoff-Bindung; und n eine ganze Zahl aus dem Bereich von 1 bis 50 ist, so daß, wenn n größer als 1 ist, jedes  $T^3$  und J unabhängig ausgewählt wird.

[0127] Die modulare Struktur  $T^2-(J-T^3)_n$  liefert einen geeigneten Zugang zu Familien von T-L-X-Verbindungen, in denen jedes Mitglied der Familie eine verschiedene T-Gruppe besitzt. Wenn T z. B.  $T^{ms}$  ist und jedes Familienmitglied wünschenswerterweise denselben MSSE besitzt, kann eine der  $T^3$ -Gruppen diese MSSE-Struktur liefern. Um für Variabilität zwischen Mitgliedern einer Familie im Hinblick auf die Masse von  $T^{ms}$  zu sorgen, kann die  $T^2$ -Gruppe unter den Familienmitgliedern variiert werden. Ein Familienmitglied kann z. B.  $T^2$  = Methyl haben, während ein anderes  $T^2$  = Ethyl hat und ein anderes  $T^2$  = Propyl hat, etc.

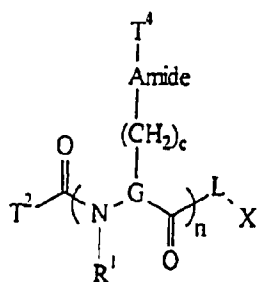
[0128] Um "grobe" oder große Sprünge in der Masse bereitzustellen, kann eine  $T^3$ -Gruppe konstruiert werden, die in signifikantem Maße (z. B. eine oder mehrere hundert) Masseeinheiten zu T-L-X hinzufügt. Solch eine  $T^3$ -Gruppe kann als eine Molekulargewichtsbereichseinstellungsgruppe ("WRA") bezeichnet werden. Eine WRA ist ziemlich nützlich, wenn man mit einem einzigen Satz von  $T^2$ -Gruppen arbeitet, die Massen haben werden, die sich über einen beschränkten Bereich erstrecken. Ein einzelner Satz von  $T^2$ -Gruppen kann verwendet

werden, um  $T^{ms}$ -Gruppen mit einem weiten Massebereich zu schaffen, einfach indem eine oder mehrere WRA- $T^3$ -Gruppen in die  $T^{ms}$  einbezogen werden. Als ein einfaches Beispiel liefert somit, wenn ein Satz von  $T^2$ -Gruppen einen Massebereich von 250–340 Daltons für die  $T^{ms}$  bereitstellt, das Hinzufügen einer einzigen WRA mit einer beispielhaften Zahl von 100 Daltons, als eine  $T^3$ -Gruppe Zugang zum Massebereich 350–440 Daltons, während derselbe Satz von  $T^2$ -Gruppen verwendet wird. In ähnlicher Weise stellt das Hinzufügen von zwei MWA-Gruppen mit 100 Daltons (jede als eine  $T^3$ -Gruppe) Zugang zum Massebereich von 450–540 Daltons bereit, wobei dieses inkrementale Hinzufügen von WRA-Gruppen fortgesetzt werden kann, um Zugang zu einem sehr großen Massebereich für die  $T^{ms}$ -Gruppe zu liefern. Bevorzugte Verbindungen der Formel  $T^2-(J-T^3)_n-L-X$  haben die Formel  $R_{VWC}-(R_{WRA})_n-R_{MSSE}-L-X$ , in der VWC eine " $T^2$ "-Gruppe ist und jede der WRA- und MSSE-Gruppen " $T^3$ "-Gruppen sind. Diese Struktur ist in **Fig. 12** veranschaulicht und stellt einen modularen Ansatz für die Herstellung von  $T^{ms}$  dar.

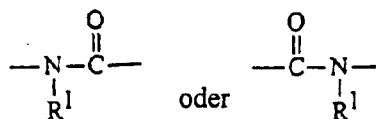
[0129] In der Formel  $T^2-(J-T^3)_n-$  sind  $T^2$  und  $T^3$  vorzugsweise ausgewählt aus Hydrocarbyl, Hydrocarbyl-O-hydrocarbylen, Hydrocarbyl-S-hydrocarbylen, Hydrocarbyl-NH-hydrocarbylen, Hydrocarbyl-amid-hydrocarbylen, N-(Hydrocarbyl hydrocarbylen, N,N-Di(hydrocarbyl)hydrocarbylen, Hydrocarbylacylhydrocarbylen, Heterocyclylhydrocarbyl, wobei das (die) Heteroatom(e) ausgewählt ist (sind) aus Sauerstoff, Stickstoff, Schwefel und Phosphor, substituiertem Heterocyclylhydrocarbyl, wobei das (die) Heteroatom(e) ausgewählt ist (sind) aus Sauerstoff, Stickstoff, Schwefel und Phosphor und die Substituenten ausgewählt sind aus Hydrocarbyl, Hydrocarbyl-O-hydrocarbylen, Hydrocarbyl-NH-hydrocarbylen, Hydrocarbyl-S-hydrocarbylen, N-(Hydrocarbyl)hydrocarbylen, N,N-Di(hydrocarbyl)hydrocarbylen und Hydrocarbylacyl-hydrocarbylen. Zusätzlich kann  $T^2$  und/oder  $T^3$  ein Derivat irgendeiner der zuvor aufgelisteten potentiellen  $T^2/T^3$ -Gruppen sein, wobei ein oder mehrere Wasserstoffe durch Fluoride ersetzt sind.

[0130] Ebenfalls im Hinblick auf die Formel  $T^2-(J-T^3)_n-$  hat ein bevorzugtes  $T^3$  die Formel  $-G(R^2)-$ , in der G eine  $C_{1-6}$ -Alkylkette mit einem einzelnen  $R^2$ -Substituenten ist. So kann, wenn G Ethylen ( $-CH_2-CH_2-$ ) ist, einer der zwei Ethylen-Kohlenstoffe einen  $R^2$ -Substituenten tragen und  $R^2$  ist ausgewählt aus Alkyl, Alkenyl, Alkynyl, Cycloalkyl, Aryl-kondensiertem Cycloalkyl, Cycloalkenyl, Aryl, Aralkyl, Aryl-substituiertem Alkenyl oder Alkynyl, Cycloalkyl-substituiertem Alkyl, Cycloalkenyl-substituiertem Cycloalkyl, Biaryl, Alkoxy, Alkenoxy, Alkinoxy, Aralkoxy, Aryl-substituiertem Alkenoxy oder Alkinoxy, Alkylamino, Alkenylamino oder Alkynylamino, Aryl-substituiertem Alkylamino, Aryl-substituiertem Alkenylamino oder Alkynylamino, Aryloxy, Arylamino, N-Alkylharnstoff-substituiertem Alkyl, N-Arylharnstoff-substituiertem Alkyl, Alkylcarbylamino-substituiertem Alkyl, Aminocarbonyl-substituiertem Alkyl, Heterocyclyl, Heterocyclyl-substituiertem Alkyl, Heterocyclyl-substituiertem Amino, Carboxyalkyl-substituiertem Aralkyl, Oxocarbocyclyl-kondensiertem Aryl und Heterocyclylalkyl; Cycloalkenyl, Aryl-substituiertem Alkyl und Aralkyl, Hydroxy-substituiertem Alkyl, Alkoxy-substituiertem Alkyl, Aralkoxy-substituiertem Alkyl, Alkoxy-substituiertem Alkyl, Aralkoxy-substituiertem Alkyl, Alkoxy-substituiertem Alkyl, Amino-substituiertem Alkyl, (Aryl-substituiertem Alkylloxycarbonylamino)-substituiertem Alkyl, Thiol-substituiertem, Alkylsulfonyl-substituiertem Alkyl, (Hydroxy-substituiertem Alkylthio)-substituiertem Alkyl, Thioalkoxy-substituiertem Alkyl, Hydrocarbylacylamino-substituiertem Alkyl, Heterocyclylocylamino-substituiertem Alkyl, Hydrocarbyl-substituiertem-heterocyclylacylamino-substituiertem Alkyl, Alkylsulfonylamino-substituiertem Alkyl, Arylsulfonylamino-substituiertem Alkyl, Morpholinoalkyl, Thiomorpholinoalkyl, Morpholinocarbonyl-substituiertem Alkyl, Thiomorpholinocarbonyl-substituiertem Alkyl, [N-(Alkyl, Alkenyl oder Alkynyl)- oder N,N-[Dialkyl, Dialkenyl, Dialkynyl oder (Alkyl, Alkenyl)-amino]carbonyl-substituiertem Alkyl, Heterocyclylaminocarbonyl, Heterocyclylalkylenaminocarbonyl, Heterocyclylaminocarbonyl-substituiertem Alkyl, Heterocyclylalkylenaminocarbonyl-substituiertem Alkyl, N,N-[Dialkyl]alkylenaminocarbonyl, N,N-[Dialkyl]alkylenaminocarbonyl-substituiertem Alkyl, Alkyl-substituiertem Heterocyclylcarbonyl, Alkyl-substituiertem Heterocyclylcarbonylalkyl, Carboxyl-substituiertem Alkyl, Dialkylamino-substituiertem Acylaminoalkyl und Aminosäureseitenketten, ausgewählt aus Arginin, Asparagin, Glutamin, S-Methylcystein, Methionin und entsprechenden Sulfoxid- und Sulfonderivaten davon, Glycin, Leucin, Isoleucin, allo-Isoleucin, tert.-Leucin, Norleucin, Phenylalanin, Tyrosin, Tryptophan, Prolin, Alanin, Ornithin, Histidin, Glutamin, Valin, Threonin, Serin, Asparaginsäure, beta-Cyanoalanin und allo-Threonin; Alinyl und Heterocyclylcarbonyl, Aminocarbonyl, Amido, Mono- oder Dialkylaminocarbonyl, Mono- oder Diarylaminocarbonyl, Alkylarylaminocarbonyl, Diarylaminocarbonyl, Mono- oder Diacylaminocarbonyl, aromatischem oder aliphatischem Acyl, Alkyl, fakultativ substituiert mit Substituenten, die ausgewählt sind aus Amino, Carboxy, Hydroxy, Mercapto, Mono- oder Dialkylamino, Mono- oder Diarylamino, Alkylarylamino, Diarylamino, Mono- oder Diacylamino, Alkoxy, Alkenoxy, Aryloxy, Thioalkoxy, Thioalkenoxo, Thioalkinoxy, Thioaryloxy und Heterocyclyl.

[0131] Eine bevorzugte Verbindung der Formel  $T^2-(J-T^3)_n-L-X$  hat die Struktur:

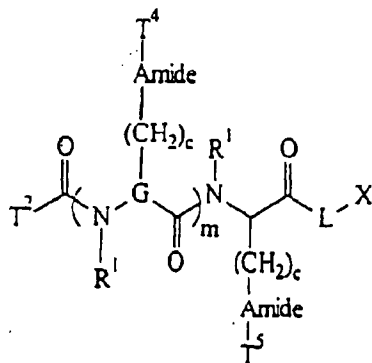


in der  $G(CH_2)_{1-6}$  ist, so daß ein Sauerstoff an einer und nur einer der  $CH_2$ -Gruppen, die durch ein einziges "G" dargestellt werden, ersetzt ist durch  $-(CH_2)_c$ -Amid- $T^4$ ;  $T^2$  und  $T^4$  organische Einheiten der Formel  $C_{2-25}N_{0-9}O_{0-9}H_\alpha F_\beta$  sind, so daß die Summe von  $\alpha$  und  $\beta$  ausreichend ist, um die sonst ungesättigten Valenzen der C-, N- und O-Atome abzusättigen; Amid



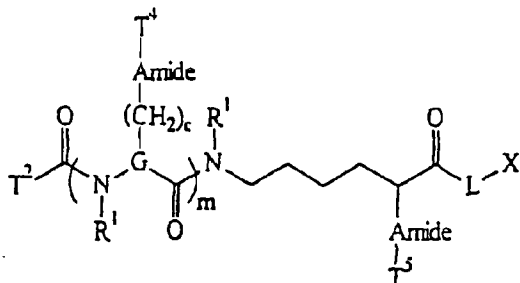
ist;  $R^1$  Wasserstoff oder  $C_{1-10}$ -Alkyl ist;  $c$  eine ganze Zahl aus dem Bereich von 0 bis 4 ist; und  $n$  eine ganze Zahl aus dem Bereich von 1 bis 50 ist, so daß, wenn  $n$  größer als 1 ist,  $G$ ,  $c$ , Amid,  $R^1$  und  $T^4$  unabhängig ausgewählt werden.

[0132] In einer weiter bevorzugten Ausführungsform hat eine Verbindung der Formel  $T^2-(J-T^3)_n-L-X$  die Struktur:



in der  $T^5$  eine organische Einheit der Formel  $C_{1-25}N_{0-9}O_{0-9}H_\alpha F_\beta$  ist, so daß die Summe von  $\alpha$  und  $\beta$  ausreichend ist, um die sonst ungesättigten Valenzen der C-, N- und O-Atome abzusättigen; und  $T^5$  ein tertiäres oder quartäres Amin oder eine organische Säure einschließt;  $m$  eine ganze Zahl aus dem Bereich von 0-49 ist und  $T^2$ ,  $T^4$ ,  $R^1$ ,  $L$  und  $X$  zuvor definiert worden sind.

[0133] Eine weitere bevorzugte Verbindung mit der Formel  $T^2-(J-T^3)_n-L-X$  hat die besondere Struktur:

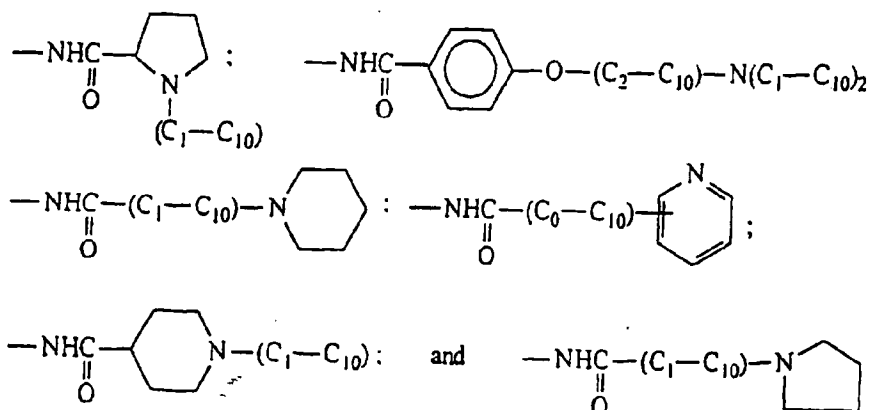


in der  $T^5$  eine organische Einheit der Formel  $C_{1-25}N_{0-9}O_{0-9}H_\alpha F_\beta$  ist, so daß die Summe von  $\alpha$  und  $\beta$  ausreichend ist, um die sonst ungesättigten Valenzen der C-, N- und O-Atome abzusättigen; und  $T^5$  ein tertiäres oder quartäres Amin oder eine organische Säure einschließt;  $m$  eine ganze Zahl aus dem Bereich von 0-49 ist und  $T^2$ ,  $T^4$ ,  $c$ ,  $R^1$ , "Amid",  $L$  und  $X$  zuvor definiert worden sind.

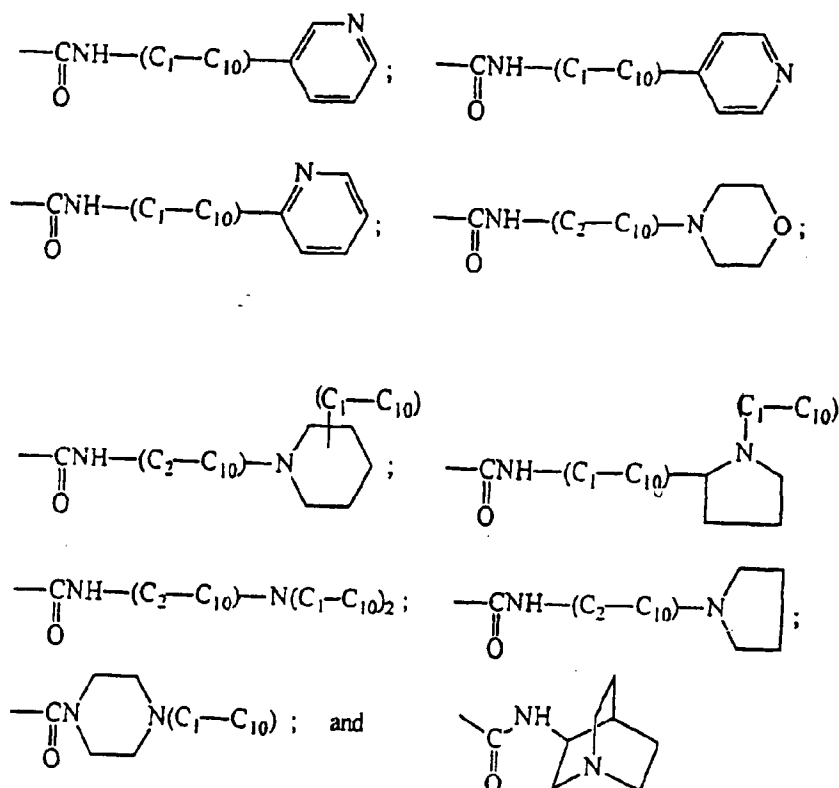
[0134] In den obigen Strukturen, die eine  $T^5$ -Gruppe besitzen, ist -Amid- $T^5$  vorzugsweise eine der folgenden Gruppen, die geeigneterweise hergestellt werden, indem organische Säuren mit freien Aminogruppen umge-



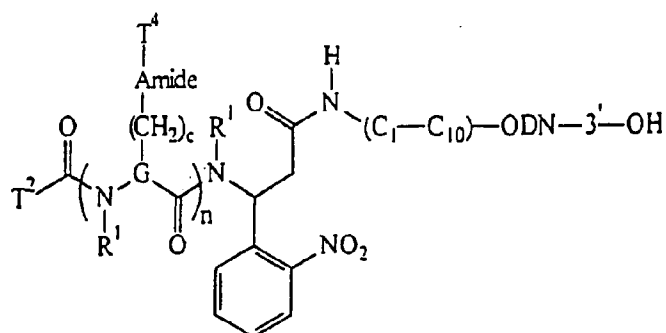
setzt werden, die an "G" gebunden sind:



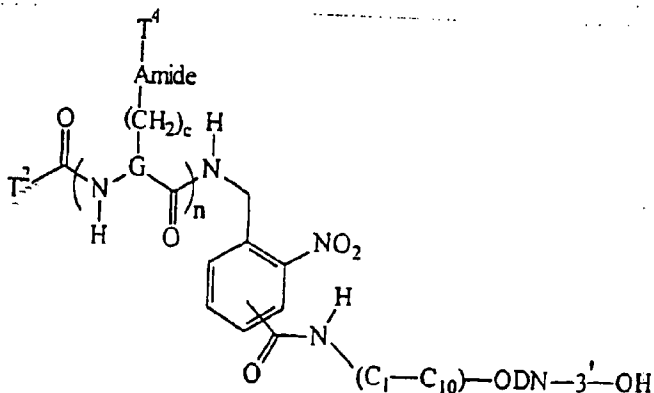
Wenn die obigen Verbindungen eine T<sup>5</sup>-Gruppe besitzen und die "G"-Gruppe eine freie Carboxylgruppe besitzt (oder ein reaktives Äquivalent davon), dann sind die folgenden Gruppen die bevorzugte -Amid-T<sup>5</sup>-Gruppen, die geeigneterweise hergestellt werden können, indem das geeignete organische Amin mit einer freien Carboxylgruppe umgesetzt wird, die an eine "G"-Gruppe gebunden ist:



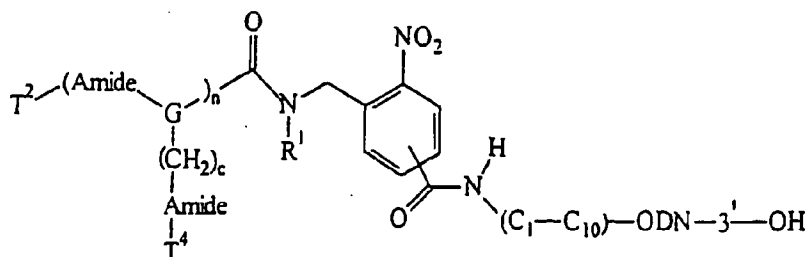
[0135] In drei bevorzugten Ausführungsformen der Erfindung hat T-L-MOI die Struktur:



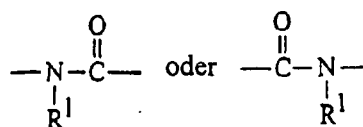
oder die Struktur:



oder die Struktur:



in denen  $T^2$  und  $T^4$  organische Einheiten der Formel  $C_{1-25}N_{0-9}O_{0-9}S_{0-3}P_{0-3}H_cF_{\beta}I_{\delta}$  sind, so daß die Summe von  $\alpha$ ,  $\beta$  und  $\delta$  ausreichend ist, um die sonst ungesättigten Valenzen der C-, N-, O-, S- und P-Atome abzusättigen;  $G(CH_2)_{1-6}$  ist, wobei ein und nur ein Wasserstoff auf den  $CH_2$ -Gruppen, die durch jedes G dargestellt werden, durch  $-(CH_2)_c$ -Amid- $T^4$  ersetzt ist; Amid



ist;  $R^1$  Wasserstoff oder  $C_{1-10}$ -Alkyl ist; c eine ganze Zahl aus dem Bereich von 0 bis 4 ist; " $C_2-C_{10}$ " eine Hydrocarbylengruppe mit von 2 bis 10 Kohlenstoffatomen darstellt, "ODN-3'-OH" ein Nukleinsäurefragment mit einer endständigen 3'-Hydroxylgruppe darstellt (d. h. ein Nukleinsäurefragment, das an  $(C_1-C_{10})$  am anderen als dem 3'-Ende des Nukleinsäurefragments gebunden ist); und n eine ganze Zahl aus dem Bereich von 1 bis 50 ist, so daß, wenn n größer als 1 ist, dann G, c, Amid,  $R^1$  und  $T^4$  unabhängig ausgewählt werden. Vorzugsweise sind nicht drei Heteroatome an ein einziges Kohlenstoffatom gebunden.

[0136] In Strukturen, wie oben angegeben, die eine  $T^2-C(=O)-N(R^1)$ -Gruppe enthalten, kann diese Gruppe dadurch gebildet werden, daß ein Amin der Formel  $HN(R^1)-$  mit einer organischen Säure umgesetzt wird, die aus den folgenden ausgewählt ist, die nur beispielhaft sind und keine erschöpfende Liste möglicher organischer Säuren darstellen: Ameisensäure, Essigsäure, Propionsäure, Fluoressigsäure, 2-Butinsäure, Cyclopropancarbonsäure, Buttersäure, Methoxyessigsäure, Difluoressigsäure, 4-Pentinsäure, Cyclo-

butancarbonsäure, 3,3-Dimethylacrylsäure, Valeriansäure, N,N-Dimethylglycin, N-Formyl-Gly-OH, Ethoxyessigsäure, (Methylthio)essigsäure, Pyrrol-2-carbonsäure, 3-Furancarbonsäure, Isoxazol-5-carbonsäure, trans-3-Hexensäure, Trifluoressigsäure, Hexensäure, Ac-Gly-OH, 2-Hydroxy-2-methylbuttersäure, Benzoesäure, Nicotinsäure, 2-Pyrazincarbonsäure, 1-Methyl-2-pyrrolcarbonsäure, 2-Cyclopenten-1-essigsäure, Cyclopentylessigsäure, (S)-(-)-2-Pyrrolidon-5-carbonsäure, N-Methyl-L-prolin, Heptensäure, Ac-b-Ala-OH, 2-Ethyl-2-hydroxybuttersäure, 2-(2-Methoxyethoxy)essigsäure, p-Tolylsäure, 6-Methylnicotinsäure, 5-Methyl-2-pyrazincarbonsäure, 2,5-Dimethylpyrrol-3-carbonsäure, 4-Fluorbenzoesäure, 3,5-Dimethylisoxazol-4-carbonsäure, 3-Cyclopentylpropionsäure, Octensäure, N,N-Dimethylsuccinamidsäure, Phenylpropionalsäure, Zimtsäure, 4-Ethylbenzoesäure, p-Anissäure; 1,2,5-Trimethylpyrrol-3-carbonsäure, 3-Fluor-4-methylbenzoesäure, Ac-DL-Propargylglycin, 3-(Trifluormethyl)buttersäure, 1-Piperidinpropionsäure, N-Acetylprolin, 3,5-Difluorbenzoesäure, Ac-L-Val-OH, Indol-2-carbonsäure, 2-Benzofurancarbonsäure, Benzotriazol-5-carbonsäure, 4-n-Propylbenzoesäure, 3-Dimethylaminobenzoessäure, 4-Ethoxybenzoesäure, 4-(Methylthio)benzoesäure, N-(2-Furoyl)glycin, 2-(Methylthio)nicotinsäure, 3-Fluor-4-methoxybenzoesäure, Tfa-Gly-OH, 2-Naphthoesäure, Chinaldinsäure, Ac-L-Ile-OH, 3-Methylinden-2-carbonsäure, 2-Chinoxalincarbonsäure, 1-Methylindol-2-carbonsäure, 2,3,6-Trifluorbenzoesäure, N-Formyl-L-Met-OH, 2-[2-(2-Methoxyethoxy)ethoxy]essigsäure, 4-n-Butylbenzoesäure, N-Benzoylglycin, 5-Fluorindol-2-carbonsäure, 4-n-Propoxybenzoesäure, 4-Acetyl-3,5-dimethyl-2-pyrrolcarbonsäure, 3,5-Dimethoxybenzoesäure, 2,6-Dimethoxynicotinsäure, Cyclohexanpentansäure, 2-Naphthylessigsäure, 4-(1H-Pyrrol-1-yl)benzoesäure, Indol-3-propionsäure, m-Trifluormethylbenzoesäure, 5-Methoxyindol-2-carbonsäure, 4-Pentylbenzoesäure, Bz-b-Ala-OH, 4-Diethylaminobenzoessäure, 4-n-Butoxybenzoesäure, 3-Methyl-5-CF<sub>3</sub>-isoxazol-4-carbonsäure, (3,4-Dimethoxyphenyl)essigsäure, 4-Biphenylcarbonsäure, Pivaloyl-Pro-OH, Octanoyl-Gly-OH, (2-Naphthoxy)essigsäure, Indol-3-buttersäure, 4-(Trifluormethyl)phenylessigsäure, 5-Methoxyindol-3-essigsäure, 4-(Trifluormethoxy)benzoesäure, Ac-L-Phe-OH, 4-Pentyloxybenzoesäure, Z-Gly-OH, 4-Carboxy-N-(fur-2-ylmethyl)pyrrolidin-2-on, 3,4-Diethoxybenzoesäure, 2,4-Dimethyl-5-CO<sub>2</sub>Et-pyrrol-3-carbonsäure, N-(2-Fluorphenyl)succinamidsäure, 3,4,5-Trimethoxybenzoesäure, N-Phenylanthranilsäure, 3-Phenoxybenzoesäure, Nonanoyl-Gly-OH, 2-Phenoxypyridin-3-carbonsäure, 2,5-Dimethyl-l-phenylpyrrol-3-carbonsäure, trans-4-(Trifluormethyl)zimtsäure, (5-Methyl-2-phenyloxazol-4-yl)essigsäure, 4-(2-Cyclohexenyl)benzoesäure, 5-Methoxy-2-methylindol-3-essigsäure, trans-4-Cotinincarbonsäure, Bz-5-Aminovaleriansäure, 4-Hexyloxybenzoesäure, N-(3-Methoxyphenyl)succinamidsäure, Z-Sar-OH, 4-(3,4-Dimethoxyphenyl)buttersäure, Ac-o-Fluor-DL-Phe-OH, N-(4-Fluorphenyl)glutaminsäure, 4'-Ethyl-4-biphenylcarbonsäure, 1,2,3,4-Tetrahydroacridincarbonsäure, 3-Phenoxyphenylessigsäure, N-(2,4-Difluorphenyl)succinamidsäure, N-Decanoyl-Gly-OH, (+)-6-Methoxy- $\alpha$ -methyl-2-naphthalinessigsäure, 3-(Trifluormethoxy)zimtsäure, N-Formyl-DL-Trp-OH, (R)-(+)- $\alpha$ -Methoxy- $\alpha$ -(trifluormethyl)phenylessigsäure, Bz-DL-Leu-OH, 4-(Trifluormethoxy)phenoxyessigsäure, 4-Heptyloxybenzoesäure, 2,3,4-Trimethoxyzimtsäure, 2,6-Dimethoxybenzoyl-Gly-OH, 3-(3,4,5-Trimethoxyphenyl)propionsäure, 2,3,4,5,6-Pentafluorphenoxyessigsäure, N-(2,4-Difluorphenyl)glutaminsäure, N-Undecanoyl-Gly-OH, 2-(4-Fluorbenzoyl)benzoesäure, 5-Trifluormethoxyindol-2-carbonsäure, N-(2,4-Difluorphenyl)diglycolamidsäure, Ac-L-Trp-OH, Tfa-L-Phenylglycin-OH, 3-Iodbenzoesäure, 3-(4-n-Pentylbenzoyl)propionsäure, 2-Phenyl-4-chinolin-carbonsäure, 4-Octyloxybenzoesäure, Bz-L-Met-OH, 3,4,5-Triethoxybenzoesäure, N-Lawoyl-Gly-OH, 3,5-Bis(trifluormethyl)benzoesäure, Ac-5-Methyl-DL-Trp-OH, 2-Iodphenylessigsäure, 3-Iod-4-methylbenzoesäure, 3-(4-n-Hexylbenzoyl)propionsäure, N-Hexanoyl-L-Phe-OH, 4-Nonyloxybenzoesäure, 4'-(Trifluormethyl)-2-biphenylcarbonsäure, Bz-L-Phe-OH, N-Tridecanoyl-Gly-OH, 3,5-Bis(trifluormethyl)phenylessigsäure, 3-(4-n-Heptylbenzoyl)propionsäure, N-Heptanoyl-L-Phe-OH, 4-Decyloxybenzoesäure, N-( $\alpha,\alpha,\alpha$ -Trifluor-m-tolyl)anthranilsäure, Nifluminsäure, 4-(2-Hydroxyhexafluorisopropyl)benzoesäure, N-Myristoyl-Gly-OH, 3-(4-n-Octylbenzoyl)propionsäure, N-Octanoyl-L-Phe-OH, 4-Undecyloxybenzoesäure, 3-(3,4,5-Trimethoxyphenyl)propionyl-Gly-OH, 8-Iodnaphthoesäure, N-Pentadecanoyl-Gly-OH, 4-Dodecyloxybenzoesäure, N-Palmitoyl-Gly-OH und N-Stearoyl-Gly-OH. Diese organischen Säuren sind von einer der folgenden Firmen erhältlich: Advanced ChemTech, Louisville, KY; Bachem Bioscience Inc., Tonance, CA; Calbiochem-Novabiochem Corp., San Diego, CA; Farchan Laboratories Inc., Gainesville FL; Lancaster Synthesis, Windham, NH; und MayBridge Chemical Company (c/o Ryan Scientific), Columbia, SC. Die Kataloge dieser Firmen verwenden die Abkürzungen, die oben verwendet werden, um die Säuren zu identifizieren.

#### f. Kombinatorische Chemie als ein Mittel zur Herstellung von Markierungen

[0137] Kombinatorische Chemie ist eine Art von Synthesestrategie, die zur Herstellung großer chemischer Bibliotheken führt (siehe z. B. PCT-Anmeldung Veröffentlichungs-Nr. WO 94/08051). Diese kombinatorischen Bibliotheken können als Markierungen zur Identifizierung von interessierenden Molekülen (MOIs) verwendet werden. Kombinatorische Chemie kann definiert werden als die systematische und wiederholte, kovalente Verknüpfung eines Satzes unterschiedlicher "Bausteine" variierender Struktur miteinander, um eine große Anzahl verschiedener Moleküleinheiten zu liefern. Bausteine können nehmen viele Formen annehmen, sowohl natürlich vorkommende als auch synthetische, wie etwa Nukleophile, Elektrophile, Diene, Alkylierungs- oder Acylierungs-

mittel, Diamine, Nukleotide, Aminosäuren, Zucker, Lipide, organische Monomere, Synthons und Kombinationen der obigen. Chemische Reaktionen, die verwendet werden, um die Bausteine zu verknüpfen, können Alkylierung, Acylierung, Oxidation, Reduktion, Hydrolyse, Substitution, Elimination, Addition, Zyklisierung, Kondensation und dergleichen einschließen. Dieser Prozeß kann Bibliotheken von Verbindungen erzeugen, die oligomer, nicht-oligomer oder Kombinationen davon sind. Wenn oligomer, können die Verbindungen verzweigt, unverzweigt oder zyklisch sein. Beispiele für oligomere Strukturen, die mit kombinatorischen Verfahren hergestellt werden können, schließen Oligopeptide, Oligonukleotide, Oligosaccharide, Poly lipide, Polyester, Polyamide, Polyurethane, Polyharnstoffe, Polyether, Poly(phosphorderivate), z. B. Phosphate, Phosphonate, Phosphoramidate, Phosphonamide, Phosphite, Phosphinamide, etc., und Poly(schwefelderivate), z. B. Sulfone, Sulfonate, Sulfate, Sulfonamide, Sulfenamide, etc., ein.

[0138] Ein üblicher Typ von oligomerer kombinatorischer Bibliothek ist die kombinatorische Peptid-Bibliothek. Kürzliche Innovationen in der Peptidchemie und Molekularbiologie haben Bibliotheken ermöglicht, die aus einigen zehn bis einigen hundert von Millionen verschiedenen Peptidsequenzen bestehen, die hergestellt und verwendet werden. Solche Bibliotheken können in drei große Kategorien unterteilt werden. Eine Kategorie von Bibliotheken umfaßt die chemische Synthese von löslichen, nicht an einen Träger gebundenen Peptidbibliotheken (z. B. Houghten et al., *Nature* 354: 84, 1991). Eine zweite Kategorie umfaßt die chemische Synthese von trägergebundenen Peptidbibliotheken, präsentiert auf festen Trägern, wie etwa Kunststoffnadeln, Harzperlen oder Baumwolle (Geysen et al., *Mol. Immunol.* 23: 709, 1986; Lam et al., *Nature* 354: 82, 1991; Eichler und Houghten, *Biochemistry* 32: 11035, 1993). In diesen ersten zwei Kategorien sind die Bausteine typischerweise L-Aminosäuren, D-Aminosäuren, nicht-natürliche Aminosäuren oder irgendeine Mischung oder Kombination davon. Eine dritte Kategorie verwendet molekularbiologische Ansätze, um Peptide oder Proteine auf der Oberfläche filamentöser Phagenteilchen oder Plasmide herzustellen (Scott und Craig, *Curr. Opin. Biotech.* 5: 40, 1994). Lösliche, nicht an Träger gebundene Peptidbibliotheken scheinen für eine Reihe von Anwendungen geeignet zu sein, einschließlich Verwendung als Markierungen. Das verfügbare Repertoire von chemischen Diversitäten in Peptidbibliotheken kann durch solche Schritte wie Permethylierung ausgedehnt werden (Ostresh et al., *Proc. Natl. Acad. Sci., USA* 91: 11138, 1994).

[0139] Zahlreiche Varianten von kombinatorischen Peptidbibliotheken sind möglich, in denen die Peptid-Rückgratkette modifiziert wird und/oder die Amidbindungen durch mimetische Gruppen ersetzt worden sind. Amid-mimetische Gruppen, die verwendet werden können, schließen Harnstoffe, Urethane und Carbonylmethylengruppen ein. Die Umstrukturierung der Rückgratkette, so daß Seitenketten von den Amid-Stickstoffen jeder Aminosäure statt von den Alpha-Kohlenstoffatomen ausgehen, ergeben Bibliotheken von Verbindungen, die als Peptoide bekannt sind (Simon et al., *Proc. Natl. Acad. Sci., USA* 89: 9367, 1992).

[0140] Ein weiterer üblicher Typ von oligomerer kombinatorischer Bibliothek ist die kombinatorische Oligonukleotidbibliothek, in der die Bausteine irgendeine Form von natürlich vorkommenden oder nicht-natürlichen Nukleotid- oder Polysaccharid-Derivaten sind, einschließlich die Fälle, in denen verschiedene organische und anorganische Gruppen die Phosphatverknüpfung ersetzen können und Stickstoff oder Schwefel Sauerstoff in einer Etherverknüpfung ersetzen können (Schneider et al., *Biochem.* 34: 9599, 1995; Freier et al., *J. Med. Chem.* 38: 344, 1995; Frank, *J. Biotechnology* 41: 259, 1995; Schneider et al., veröffentlichte PCT WO 942052; Ecker et al., *Nucleic Acids Res.* 21: 1853, 1993).

[0141] Vor kurzem ist die kombinatorische Herstellung von Sammlungen nicht-oligomerer, kleiner Molekülverbindungen beschrieben worden (DeWitt et al., *Proc. Natl. Acad. Sci. USA* 90: 690, 1993; Bunin et al., *Proc. Natl. Acad. Sci., USA* 91: 4708, 1994). Strukturen, die für die Einstellung von Kleinmolekül-Bibliotheken geeignet sind, umfassen eine breite Vielfalt organischer Moleküle, z. B. Heterozyklen, Aromaten, Alizyklen, Aliphaten, Steroide, Antibiotika, Enzyminhibitoren, Liganden, Hormone, pharmazeutische Wirkstoffe, Alkaloide, Opioide, Terpene, Porphyrine, Toxine, Katalysatoren sowie Kombinationen davon.

#### g. Spezifische Verfahren zur kombinatorischen Synthese von Markierungen

[0142] Zwei Verfahren zur Herstellung und Verwendung eines unterschiedlichen Satzes Amin-enthaltender MS-Markierungen sind unten umrissen. In beiden Verfahren wird Festphasensynthese eingesetzt, um simultane parallele Synthese einer großen Anzahl markierter Linker zu ermöglichen, unter Verwendung der Techniken der kombinatorischen Chemie. Beim ersten Verfahren führt die letztendliche Abspaltung der Markierung vom Oligonukleotid zur Freisetzung eines Carbonsäureamids. Beim zweiten Verfahren erzeugt die Abspaltung der Markierung eine Carbonsäure. Die chemischen Komponenten und Verknüpfungselemente, die in diesen Verfahren verwendet werden, werden wie folgt abgekürzt:

R	Harz
FMOC	Fluorenylmethoxycarbonyl-Schutzgruppe
All	Allyl-Schutzgruppe
CO <sub>2</sub> H	Carbonsäuregruppe
CONH <sub>2</sub>	Carbonsäureamidgruppe
NH <sub>2</sub>	Aminogruppe
OH	Hydroxylgruppe
CONH	Amidverknüpfung
COO	Esterverknüpfung
NH <sub>2</sub> -Rink-CO <sub>2</sub> H	4-[(α-Amino)-2,4-dimethoxybenzyl]-phenoxybuttersäure( Rink-Linker)
OH-1MeO-CO <sub>2</sub> H	(4-Hydroxymethyl)phenoxybuttersäure
OH-2MeO-CO <sub>2</sub> H	(4-Hydroxymethyl-3-methoxy)phenoxyessigsäure
NH <sub>2</sub> -A-COOH	Aminosäure mit aliphatischer oder aromatischer Aminfunktionalität in der Seitenkette
X1 . . . Xn-COOH	Satz von n verschiedenen Carbonsäuren mit einzigartigen relativen Molekülmassen
Oligo1 . . . Oligo(n)	Satz von n Oligonukleotiden
HBTU	O-Benzotriazol-1-yl-N,N,N',N'-tetramethyluroniumhexafluorophosphat

[0143] Die Abfolge der Schritte in Verfahren 1 ist wie folgt:

OH-2MeO-CONH-R

↓ FMOC-NH-Rink-CO<sub>2</sub>H; Kopplung (z.B. HBTU)

FMOC-NH-Rink-COO-2MeO-CONH-R

↓ Piperidin (entferne FMOC)

NH<sub>2</sub>-Rink-COO-2MeO-CONH-R

↓ FMOC-NH-A-COOH; Kopplung (z.B. HBTU)

FMOC-NH-A-CONH-Rink-COO-2MeO-CONH-R

↓ Piperidin (entferne FMOC)

NH<sub>2</sub>-A-CONH-Rink-COO-2MeO-CONH-R

↓ unterteile in n Aliquote

↓↓↓↓↓ kopple an n unterschiedliche Säuren X1...Xn-COOH

X1...Xn-CONH-A-CONH-Rink-COO-2MeO-CONH-R

↓↓↓↓↓ spalte markierte Linker von Harz mit 1 % TFA ab

X1...Xn-CONH-A-CONH-Rink-CO<sub>2</sub>H

↓↓↓↓↓ kopple an n Oligos (Oligo1...Oligo(n)) (z.B. über Pfp-Ester)

X1...Xn-CONH-A-CONH-Rink-CONH-Oligo1...Oligo(n)

↓ sammle markierte Oligos

↓ führe Sequenzierungsreaktion durch

↓ trenne Fragmente unterschiedlicher Länge aus

Sequenzierungsreaktion auf (z.B. über HPLC oder CE)

↓ spalte Markierungen von Linkern mit 25 %-100 % TFA ab

$X_1 \dots X_n\text{-CONH-A-CONH}$

↓

analysiere durch Massenspektrometrie

[0144] Die Abfolge der Schritte in Verfahren 2 ist wie folgt:

$\text{OH-1MeO-CO}_2\text{-All}$

↓  $\text{FMOC-NH-A-CO}_2\text{H}$ ; Kopplung (e.g. HBTU)

$\text{FMOC-NH-A-COO-1MeO-CO}_2\text{-All}$

↓ Palladium (entferne Allyl)

$\text{FMOC-NH-A-COO-1MeO-CO}_2\text{H}$

↓  $\text{OH-2MeO-CONH-R}$ ; Kopplung (z.B. HBTU)

$\text{FMOC-NH-A-COO-1MeO-COO-2MeO-CONH-R}$

↓ Piperidin (entferne FMOC)

$\text{NH}_2\text{-A-COO-1MeO-COO-2MeO-CONH-R}$

↓ unterteile in n Aliquots

↓↓↓↓↓ kopple an n unterschiedliche Säuren  $X_1 \dots X_n\text{-CO}_2\text{H}$

$X_1 \dots X_n\text{-CONH-A-COO-1MeO-COO-2MeO-CONH-R}$

↓↓↓↓↓ spalte markierte Linker von Harz mit 1 % TFA ab

$X_1 \dots X_n\text{-CONH-A-COO-1MeO-CO}_2\text{H}$

↓↓↓↓↓ kopple an n Oligos (Oligo1...Oligo(n)) (z.B. über Pfp-Ester)

$X_1 \dots X_n\text{-CONH-A-COO-1MeO-CONH-Oligo1...Oligo(n)}$

↓ sammle markierte Oligos

↓ führe Sequenzierungsreaktion durch

↓ trenne Fragmente unterschiedlicher Länge aus

Sequenzierungsreaktion auf (z.B. über HPLC oder CE)

↓ spalte Markierungen von Linkern mit 25 %-100 % TFA

$X_1 \dots X_n\text{-CONH-A-CO}_2\text{H}$

↓

analysiere durch Massenspektrometrie

h. Phosphoramidit und verwandte Verfahren zur Marker-Synthese

[0145] Festphasesynthese von natürlichen Polymeren wurde ursprünglich gleichzeitig durch Merrifield (Merrifield, 1963) und für Peptidchemie entwickelt und anschließend durch Letsinger (Letsinger und Mahadevan 1965) an die Oligonukleotidsynthese angepaßt. Das Konzept weist vier grundsätzliche Aspekte auf: I. Das Oli-

gonukleotid wird synthetisiert, während es kovalent an einen festen Träger gebunden ist; II. Überschüssige lösliche geschützte Nukleotide und Kopplungsreagenz können eine Reaktion fast bis zur Vervollständigung antreiben; III. Die Reaktion wird in einem einzelnen Reagenz durchgeführt, um mechanische Verluste aufgrund der festen Träger-Manipulation zu verringern, was eine Synthese mit geringen Mengen an Ausgangsmaterialien ermöglicht; und IV. Die heterogenen Reaktionen werden standardisiert und diese Verfahren werden leicht automatisiert.

[0146] Das am meisten verbreitete Verfahren zur Synthese von Oligonukleotiden ist der Phosphit-Triester-Ansatz. Ein weiteres jedoch weniger verbreitetes Verfahren der Synthese ist der H-Phosphonat-Ansatz.

#### Der Phosphit-Triester-Ansatz zur Oligonukleotid-Synthese

[0147] Die Entwicklung dieses Verfahrens wurde 1975 begonnen, als Letsinger (Letsinger et al. 1975) das symmetrische Phosphit-Reagenz Methoxyphosphodichloridit einführen. Obwohl die Kopplungszeiten dramatisch verringert wurden, weist diese Verbindung den Nachteil auf, daß sie zu reaktiv ist, was die Handhabung bei Raumtemperatur sehr schwierig macht und die Lagerung von Phosphitmonomeren sogar bei  $-10^{\circ}\text{C}$  unmöglich macht. Die Reaktion mit geschützten Nukleotiden führt zu der Herstellung von großen Mengen von symmetrischem 3'-3'-Dimer (Letsinger et al., 1982).

[0148] 1981 löste die Einführung des neuen phosphitylierenden Mittels N,N-Dimethylaminomethoxyphosphin (Beaucage und Caruthers 1981) nicht nur das Problem der Bildung von 3'-3'-Dimeren während der Phosphitylierung, sondern führte auch zur Erzeugung von Desoxyribonukleosid-Phosphitderivaten, die bis zu einem bestimmten Ausmaß stabil gegenüber Sauerstoff und atmosphärischer Luftfeuchtigkeit bei Raumtemperatur sind. Als die am meisten brauchbare Verbindung erwies sich N,N-Diisopropylamin (Adams et al. 1983, McBride und Caruthers 1983) das leicht auf Silicagelsäulen gereinigt werden kann und das als trockene Pulver bei Raumtemperatur stabil ist.

[0149] Die durch dieses Verfahren verwendete 5'-Schutzgruppe ist das Dimethoxytrityl (DMT). Diese Gruppe wird vollständig durch Behandlung mit Trichloressigsäure (1–3% w/v) in Dichlormethan in weniger als einer Minute abgespalten. Sobald die Schutzgruppe entfernt ist, ist das freie 5'-Hydroxyl zur Kopplung an den nächsten Nukleosidbaustein erhältlich.

[0150] Anders als die Phosphodichloridite in Letsingers Ansatz und das Phosphomonochloridit/Tetrazol in der Ursprungsarbeit von Caruthers (Caruthers et al. 1980) können die Phosphoramidite nicht direkt mit einer freien 5'-Hydroxylfunktion auf einer wachsenden Kette reagieren. Sie müssen zunächst durch Behandlung mit einer schwachen Säure, wie zum Beispiel Tetrazol, aktiviert werden. Tetrazol wurde als eine zweifache Rolle spielend gezeigt (Bernier et al., 1990; Dahl et al., 1990): sie protoniert die Dialkylaminogruppe der Phosphoramiditfunktion und wirkt als nächstes als ein Nukleophil wodurch ein sehr reaktives Tetrazolphosphan-Intermediat erzeugt wird. Die Kopplungsreaktionen mit diesen Desoxynukleosid-Phosphoramiditreagenzien sind sehr schnell (weniger als 2 Minuten) und fast quantitativ.

[0151] Weil die Kopplungsreaktion in einer begrenzten Zeitdauer nicht quantitativ sein kann, wird bei jedem Kopplungsschritt ein kleiner Prozentsatz von verkürzten Sequenzen hergestellt. Diese Reaktionsfehlschläge enthalten 5'-Hydroxyle. Falls diesen falschen Sequenzen ermöglicht würde, weiter zu reagieren, wäre es schwierig, das Produkt aus dem Sequenzgemisch zu isolieren. Dieses Problem wird größtenteils durch maskieren der verbleibenden freien 5'-Hydroxyle durch Acetylierung überwunden. Dieser Maskierungsschritt wird mit dem starken Acetylierungsreagenz N-Acetyldimethylaminopyridinium erreicht, das sich bei Reaktion von äquimolaren Mengen von Essigsäureanhydrid und 4-Dimethylaminopyridin (DMAP) bildet. Die Reaktion ist in 0,5 Minuten fast quantitativ. Unter der Verwendung von N-Methylimidazol anstelle von DMAP wird das Oligonukleotid verbesserte biologische Eigenschaften aufweisen (Fenace et al., 1989).

[0152] Die neu gebildete Phosphit-Internukleotidverbindung ist instabil und sowohl gegenüber saurer als auch basischer Spaltung empfindlich. Daher wird der trivalente Phosphitester nach der Maskierung zu einem stabilen pentavalenten Phosphatriester oxidiert. Jod wird als ein mildes Oxidationsmittel in basischer Tetrahydrofwanlösung mit Wasser als dem Sauerstoff-Donor verwendet. Die Reaktionen sind extrem schnell, wobei sie in 30 Sekunden quantitativ ist. Die Oxidation vervollständigt den Nukleotidadditionszyklus. Die Kettenverlängerung kann nach Entfernen der Dimethoxytritylgruppe und dem 5'-Ende der wachsenden Kette und dem Wiederholen eines weiteren Zyklus von der Nukleotidaddition fortschreiten. Die Vervollständigung der Abspaltung von dem Träger und die gleichzeitige Phasen- und Phosphatentschätzung werden durch Behandlung mit konzentriertem Ammoniumhydroxyd erreicht.

#### H-Phosphonat-Verfahren der Oligonukleotid-Synthese

[0153] Die Verwendung eines Nukleosid-H-Phosphonats wurde zuerst durch Todd und Mitarbeiter berichtet (Hall et al. 1957). Danach blieb die H-Phosphonat-Chemie bis in die 80er Jahre unerforscht. 1985 und 1986 wurde das Wiederaufleben dieses Ansatzes durch Garegg et al. (1985, 1986a, 1986b) und Froehler et al.

(1986a und 1986b) eingeführt.

[0154] In diesem Verfahren ist das aktivierbare Monomer ein 5'-DNT-Base-geschütztes-Nukleosid-3'-Hydrogenphosphonat. In diesen Monomeren macht die Anwesenheit der H-Phosphonatgruppe den Phosphatschutz überflüssig.

[0155] Die selben Basenschutzgruppen werden verwendet, wie im Phosphat-Triester-Ansatz. Daher ist die Schutzstrategie dieselbe für die Hydroxyl- und die exocyclischen Amine auf den Heterocyclen der Entschützung.

[0156] H-Phosphonat-Synthese-Zyklus. Es gibt keinen Oxidationsschritt während der Kettenverlängerung in der Oligonukleotid-Synthese gemäß den H-Phosphonat-Verfahren. Die Oxidation wird am Ende der Synthese durchgeführt.

[0157] Der Kopplungsprozess der H-Phosphonat-Synthese wird durch ein gehindertes Acylchlorid aktiviert und das gebildete Anhydrid wird dazu verwendet, um mit einem freien Oligonukleotid 5'-Hydroxylende zu reagieren, wobei ein H-Phosphonat-Analog der Internukleotid-Brücke gebildet wird. Die Ausbeuten sind ungefähr 96–99%. Von Pivaloylchlorid und 1-Adamantancarbonylchlorid (Andrus et al., 1988) wurde berichtet, daß dies die besten Aktivatoren sind. Jedoch werden einige Nebenreaktionen zwischen den kondensierenden Reagenzien und den Ausführungsmaterial während der Kondensation beobachtet und dies führt zu abnehmenden Ausbeuten der gewünschten Verbindung. Insbesondere führte die Präaktivierung eines nukleosidischen 3'-H-Phosphonats, gefolgt von der Addition an einer OH-Komponente, die gewöhnlicherweise bei der Synthese auf Polymerträgern stattfindet, zu geringeren Ausbeuten der H-Phosphonat-Diester. Die andere Seite der Reaktion ist eine Modifikation von heterozyklischen Basen von Nukleotiden (Acylierung oder Phosphitylierung von Guanin und Thymin), während der Kondensation). Das Maskierungsreagenz der Phosphoramidit-Chemie (Essigsäureanhydrid/N-Methylimidazol) ist für den H-Phosphonat-Ansatz nicht geeignet. Cyanoethyl-H-Phosphonat (Gaffney und Jones, 1988) oder Iso-Propyl-H-Phosphonat (Andres et al., 1988), aktiviert durch Acetylchlorid oder PFPC, kann verwendet werden.

[0158] Nach Vervollständigung der Sequenz werden alle H-Phosphonat-Bindungen gleichzeitig zu Phosphodiesterverbindungen oxidiert. Anstelle der Oxidation mit Jod sind H-Phosphonat-Desoxyribonukleotide in der Lage, in DNA-Analoga zu Phosphorthioaten, Phosphoramidaten oder Phosphotriestern konvertiert zu werden (Froehler, 1986c, Froehler et al., 1988). Die Vorteile dieses Verfahrens sind eine erhöhte Monomerstabilität, die Präparation von <sup>35</sup>Smarkierten Oligonukleotiden (Stein et al., 1990) und die Möglichkeit der Wiederverwendung des Überschusses von aktiviertem Nukleosid, das nicht reagiert hat (Seliger und Rösch, 1990). H-Phosphonate sind käuflich erhältlich von Glen Research (Herndon, VA).

[0159] Die vorliegende Erfindung stellt Marker zur Verfügung die leicht in die oben beschriebenen Standardoligonukleotidsynthesen inkorporiert werden können. Die vorliegende Erfindung stellt einen CMST-Marker zur Anbringung an eine Nukleinsäure während der Festphase-Synthese zur Verfügung. Der CMST-Marker ist ein Phosphoramidit eines CMST.

[0160] Ein Verfahren gemäß der vorliegenden Erfindung schließt die Anbringung eines CMST-Markers an eine Nukleinsäure während Festphase-Synthese ein. Genauer gesagt wird ein Phosphoramidit eines CMST an ein Träger-gebundenes Oligodesoxynukleotid kondensiert. Die Anpassung von Festphase-Phosphoramidit-Chemie an CMST-Technologie vereinfacht die Reinigung und erleichtert die Automatisierung des Verfahrens, wodurch der Durchsatz erhöht wird. Dieses ermöglicht die einfache Herstellung der markierten Moleküle.

[0161] Wie oben beschrieben wurde das Phosphoramiditverfahren der Oligonukleotid-Synthese automatisiert und wird weit verwendet. Ein Phosphoramidit eines CMST-Markers stellt ein bequemes Verfahren der Markierung von Oligonukleotiden mit CMST-Markern zur Verfügung. Gemäß der vorliegenden Erfindung wird ein Phosphoramidit eines CMST-Markers an eine Oligonukleotidkette kondensiert, die an einen festen Träger gebunden wurde. Der Träger kann von jeder für die Festphase von Nukleinsäuren geeigneten Art sein. Das Polynukleosid kann ein Ribonukleosid- oder ein Desoxyribonukleosid-Phosphodiester sein. Auch können Polynukleosid-Analoga, einschließlich, jedoch nicht begrenzt auf, Phosphorthioate, Methylphosphonate, usw., verwendet werden.

[0162] Ein CMST-Phosphoramidit gemäß der vorliegenden Erfindung kann von der generellen Struktur T<sup>MS</sup>-L-X sein, wobei T<sup>MS</sup> durch Massenspektrometrie nachgewiesen wird, L ein photochemischer Linker ist und x eine Gruppe ist, die ein Koppeln an Polynukleosid an einem festen Träger ermöglicht.

[0163] Die automatisierte Synthese von Oligonukleotiden verwendete Phosphoramidit-Chemie<sup>1</sup> und eine Vielzahl von Phosphoramidit-Reagenzien (z. B. Biotin<sup>2</sup>) wurden entwickelt, die vorteilhafterweise auf dieser Chemie beruhen, um Oligonukleotide zu markieren (siehe Fig. 24)<sup>3</sup>. Die Anpassung dieser Chemie, um eine Massenmarkierung von Festphase-getragenen Oligonukleotiden zu ermöglichen, vereinfacht die Reinigung und erleichtert die Automatisierung des Verfahrens, wodurch der Durchsatz erhöht wird.

[0164] Ein vorgeschlagenes Schema, das von der bereits hergestellten Markersäure ausgeht, ist in Fig. 25 gezeigt.



## 2. Linker

[0165] Eine "Linker"-Komponente (oder L), wie hierin verwendet, bedeutet entweder eine direkte kovalente Bindung oder eine organische chemische Gruppe, die verwendet wird, um eine "Markierung" (oder T) an ein "interessierendes Molekül" (oder MOI) durch kovalente chemische Bindungen zu binden. Zusätzlich ist die direkte Bindung selbst oder eine oder mehrere Bindungen innerhalb der Linker-Komponente unter Bedingungen spaltbar, die ermöglichen, daß T vom Rest der T-L-X-Verbindung (einschließlich der MOI-Komponente) freigesetzt (mit anderen Worten abgespalten) wird. Die variable Markierungskomponente, die innerhalb von T vorhanden ist, sollte gegenüber den Spaltungsbedingungen stabil sein. Vorzugsweise kann die Spaltung schnell durchgeführt werden; innerhalb weniger Minuten und vorzugsweise innerhalb etwa 15 Sekunden oder weniger.

[0166] Im allgemeinen wird ein Linker verwendet, um jede aus einem großen Satz von Markierungen mit jedem aus einem ähnlich großen Satz von MOIs zu verknüpfen. Typischerweise wird eine einzige Markierungs-Linker-Kombination mit jedem MOI verknüpft (um verschiedene T-L-MOI zu ergeben), aber in einigen Fällen kann mehr als eine Markierungs-Linker-Kombination mit jedem individuellen MOI verknüpft werden (um verschiedene (T-L)<sub>n</sub>-MOI zu ergeben). In einer weiteren Ausführungsform der vorliegenden Erfindung werden zwei oder mehr Markierungen an einen einzigen Linker durch mehrere, unabhängige Stellen auf dem Linker gebunden, und diese Mehrfachmarkierungs-Linker-Kombination wird dann an ein einzelnes MOI gebunden (um verschiedene (T)<sub>n</sub>-L-MOI zu ergeben).

[0167] Nach verschiedenen Manipulationen des Satzes markierter MOIs werden spezielle chemische und/oder physikalische Bedingungen verwendet, um eine oder mehrere kovalente Bindungen im Linker zu spalten, was zur Freisetzung der Markierungen von den MOIs führt. Die spaltbare(n) Bindungen) kann (können) einige derselben Bindungen sein, die gebildet wurden, als die Markierung, der Linker und das MOI miteinander verknüpft wurden, oder nicht. Die Konstruktion des Linkers wird zum großen Teil die Bedingungen bestimmen, unter denen die Spaltung durchgeführt werden kann. Demgemäß können Linker durch die Spaltungsbedingungen identifiziert werden, gegenüber denen sie besonders empfindlich sind. Wenn ein Linker photolabil ist (d. h. zur Spaltung durch Einwirkung von aktinischer Strahlung neigt), kann dem Linker die Bezeichnung L<sup>hu</sup> gegeben werden. In ähnlicher Weise können die Bezeichnungen L<sup>Säure</sup>, L<sup>Base</sup>, L<sup>[O]</sup>, L<sup>[R]</sup>, L<sup>enz</sup>, L<sup>elc</sup>, L<sup>A</sup> und L<sup>ss</sup> verwendet werden, um Linker zu bezeichnen, die besonders empfindlich sind gegenüber der Spaltung durch Säure, Base, chemische Oxidation, chemische Reduktion, die katalytische Aktivität eines Enzyms (einfacher "Enzym"), elektrochemische Oxidation oder Reduktion, erhöhte Temperatur ("thermisch") bzw. Thiol austausch.

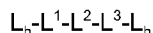
[0168] Bestimmte Linker-Typen sind gegenüber einem einzigen Typ von Spaltungsbedingung labil, wohingegen andere gegenüber mehreren Typen von Spaltungsbedingungen labil sind. Zusätzlich kann jede der Markierungs-Bindungsstellen in Linkern, die mehrere Markierungen binden können (um (T)<sub>n</sub>-L-MOI-Strukturen zu ergeben), gegenüber unterschiedlichen Spaltungsbedingungen labil sein. Bei einem Linker mit zwei daran gebundenen Markierungen kann z. B. eine der Markierungen nur gegenüber Base labil sein und die andere nur gegenüber Photolyse.

[0169] Ein Linker, der in der vorliegenden Erfindung nützlich ist, besitzt mehrere Eigenschaften:

- 1) Der Linker besitzt einen chemischen Griff (L<sub>n</sub>), durch den er mit einem MOI verknüpft werden kann.
- 2) Der Linker besitzt einen zweiten, separaten chemischen Griff (L<sub>n</sub>), durch den die Markierung an den Linker gebunden wird. Wenn mehrere Markierungen an einen einzigen Linker gebunden werden ((T)<sub>n</sub>-L-MOI-Strukturen), dann existiert für jede Markierung ein separater Griff.
- 3) Der Linker ist gegenüber allen Manipulationen stabil, denen er unterworfen wird, mit Ausnahme der Bedingungen, die eine Spaltung ermöglichen, so daß eine T-enthaltende Einheit vom Rest der Verbindung, einschließlich MOI, freigesetzt wird. Somit ist der Linker während der Bindung der Markierung an den Linker, der Bindung des Linkers an das MOI und allen Manipulationen des MOIs, während die Markierung und der Linker (T-L) daran gebunden sind, stabil.
- 4) Der Linker beeinflusst die Manipulationen, die am MOI durchgeführt werden, während das T-L daran gebunden ist, nicht signifikant. Wenn das T-L z. B. an ein Oligonukleotid gebunden ist, darf das T-L irgendwelche Hybridisierungs- oder enzymatischen Reaktionen (z. B. PCR), die am Oligonukleotid durchgeführt werden, nicht signifikant beeinflussen. In ähnlicher Weise darf, wenn das T-L an einen Antikörper gebunden ist, es nicht die Antigen-Erkennung durch den Antikörper beeinflussen.
- 5) Die Abspaltung der Markierung vom Rest der Verbindung tritt in einer in hohem Maße kontrollierten Art und Weise ein, unter Verwendung physikalischer oder chemischer Verfahren, die die Nachweisbarkeit der Markierung nicht nachteilig beeinflussen.

[0170] Für jeden gegebenen Linker ist es bevorzugt, daß der Linker mit einer breiten Vielfalt von MOIs verknüpfbar ist und daß eine breite Vielfalt von Markierungen mit dem Linker verknüpfbar ist. Eine solche Flexibilität ist vorteilhaft, weil sie erlaubt, daß eine Bibliothek von T-L-Konjugaten, wenn sie erst einmal hergestellt ist, mit verschiedenen unterschiedlichen Sätzen von MOIs verwendet werden kann.

[0171] Wie oben erläutert, hat ein bevorzugter Linker die Formel



[0172] in der jedes  $L_n$  ein reaktiver Griff ist, der verwendet werden kann, um den Linker mit einem Markierungs-Reaktanten und einem Reaktanten aus einem interessierenden Molekül zu verknüpfen.  $L^2$  ist ein wesentlicher Teil des Linkers, weil  $L^2$  dem Linker Labilität verleiht.  $L^1$  und  $L^3$  sind fakultative Gruppen, die wirkungsvoll dazu dienen,  $L^2$  von den Griffen  $L_n$  zu trennen.

[0173]  $L^1$  (das definitionsgemäß näher an T ist als  $L^3$ ) dient dazu, T von der erforderlichen labilen Einheit  $L^2$  zu trennen. Diese Trennung kann nützlich sein, wenn die Spaltungsreaktion besonders reaktive Spezies (z. B. freie Radikale) erzeugt, die zufällige Veränderungen in der Struktur der T-enthaltenden Einheit bewirken könnten. Da die Spaltungsstelle weiter von der T-enthaltenden Einheit getrennt wird, besteht eine verringerte Wahrscheinlichkeit, daß reaktive Spezies, die an der Spaltungsstelle gebildet werden, die Struktur der T-enthaltenden Einheit aufbrechen werden. Auch können, da die Atome in  $L^1$  typischerweise in der T-enthaltenden Einheit vorhanden sein werden, diese  $L^1$ -Atome der T-enthaltenden Einheit eine wünschenswerte Qualität verleihen. Wenn die T-enthaltende Einheit z. B. eine  $T^{ms}$ -enthaltende ist und ein gehindertes Amin wünschenswerterweise als Teil der Struktur der  $T^{ms}$ -enthaltenden Einheit vorhanden ist (um z.B. als ein MSSE zu dienen), kann das gehinderte Amin in der labilen  $L^1$ -Einheit vorhanden sein.

[0174] In anderen Fällen können  $L^1$  und/oder  $L^3$  in einer Linkerkomponente vorhanden sein, weil der kommerzielle Lieferant eines Linkers sich entschieden hat, den Linker in einer Form mit solch einer  $L^1$ - und/oder  $L^3$ -Gruppe zu vertreiben. In solch einem Fall besteht kein Nachteil darin, Linker mit  $L^1$ - und/oder  $L^3$ -Gruppen zu verwenden (so lange diese Gruppen die Spaltungsreaktion nicht hemmen), selbst dann, wenn sie den Verbindungen, die sie umfassen, keine besonderen Leistungsvorteile verleihen. Somit erlaubt die vorliegende Erfindung, daß  $L^1$ - und/oder  $L^3$ -Gruppen in der Linkerkomponente vorhanden sind.

[0175]  $L^1$ - und/oder  $L^3$ -Gruppen können eine direkte Bindung sein (wobei in diesem Falle die Gruppe effektiv nicht vorhanden ist), eine Hydrocarbylen-Gruppe (z. B. Alkylen, Arylen, Cycloalkylen, etc.), -O-Hydrocarbylen (z. B. -O-CH<sub>2</sub>-, O-CH<sub>2</sub>CH(CH<sub>3</sub>)-, etc.) oder Hydrocarbylen-(O-hydrocarbylen)<sub>w</sub>, wobei w eine ganze Zahl aus dem Bereich von 1 bis etwa 10 ist (z. B. -CH<sub>2</sub>-O-Ar-, -CH<sub>2</sub>-(O-CH<sub>2</sub>CH<sub>2</sub>)<sub>4</sub>-, etc.).

[0176] Mit dem Auftauchen der Festphasensynthese ist eine große Menge an Literatur zu Linkern entstanden, die gegenüber spezifischen Reaktionsbedingungen labil sind. Bei der typischen Festphasensynthese wird ein fester Träger durch einen labilen Linker an eine reaktive Stelle gebunden, und ein zu synthetisierendes Molekül wird an der reaktiven Stelle erzeugt. Wenn das Molekül vollständig synthetisiert ist, wird das Konstrukt aus festem Träger, Linker und Molekül Spaltungsbedingungen unterworfen, die das Molekül vom festen Träger freisetzen. Die labilen Linker, die zur Verwendung in diesem Zusammenhang entwickelt worden sind (oder die in diesem Zusammenhang verwendet werden können), können auch ohne weiteres als der Linker-Reaktant in der vorliegenden Erfindung verwendet werden.

[0177] Lloyd-Williams, P., et al., "Convergent Solid-Phase Peptide Synthesis", Tetrahedron Report Nr. 347, 49(48): 11065-11133 (1993) liefert eine ausführliche Diskussion von Linkern, die gegenüber aktinischer Strahlung (d. h. Photolyse) sowie sauren, basischen und anderen Spaltungsbedingungen labil sind. Zusätzliche Informationsquellen über labile Linker sind leicht erhältlich.

[0178] Wie oben beschrieben, werden unterschiedliche Linkerkonstruktionen unter unterschiedlichen spezifischen physikalischen oder chemischen Bedingungen Spaltbarkeit ("Labilität") verleihen. Beispiele für Bedingungen, die dazu dienen, verschiedene Linkerkonstruktionen zu spalten, schließen saure, basische, Oxidations-, Reduktions-, Fluorid-, Thiolaustausch-, Photolyse- und enzymatische Bedingungen ein.

[0179] Beispiele für spaltbare Linker, die die allgemeinen Kriterien für Linker erfüllen, die oben aufgelistet sind, werden den Fachleuten gut bekannt sein und schließen diejenigen ein, die im Katalog zu finden sind, der von Pierce (Rockford, IL) erhältlich ist. Beispiele schließen ein:

- Ethylenglycobis(succinimidylsuccinat) (EGS), ein aminreaktives Vernetzungsreagens, das mit Hydroxylamin spaltbar ist (1 M bei 37°C für 3-6 Stunden);
- Disuccinimidyltartrat (DST) und Sulfo-DST, die aminreaktive Vernetzungsreagentien sind, spaltbar mit 0,015 M Natriumperiodat;
- Bis[2-(succinimidylloxycarbonyloxy)ethyl]sulfon (BSOCOES) und Sulfo-BSOCOES, die aminreaktive Vernetzungsmittel sind, spaltbar durch Base (pH 11,6);
- 1,4-Di[3'-(2'-pyridyldithio(propionamido))butan (DPDPB), ein Pyridyldithiol-Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist;
- N-[4-(p-Azidosalicylamido)-butyl]-3'-(2'-pyridyldithio)propionamid (APDP), ein Pyridyldithiol-Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist;
- Bis-[beta-4-(azidosalicylamido)ethyl]-disulfid, ein photoreaktiver Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist;
- N-Succinimidyl-(4-azidophenyl)-1,3'-dithiopropionat (SADP), ein photoreaktiver Vernetzer, der durch Thio-

laustausch oder Reduktion spaltbar ist;

- Sulfosuccinimidyl-2-(7-azido-4-methylcumarin-3-acetamid)ethyl-1,3'-dithiopropionat (SAED), ein photo-reaktiver Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist;
- Sulfosuccinimidyl-2-(m-azido-o-nitrobenzamido)-ethyl-1,3'-dithiopropionat (SAND), ein photoreaktiver Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist.

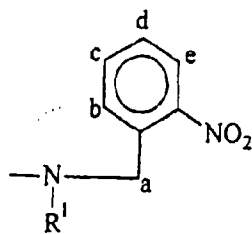
[0180] Andere Beispiele für spaltbare Linker und die Spaltungsbedingungen, die verwendet werden können, um Markierungen freizusetzen, sind wie folgt. Eine Silyl-Verknüpfungsgruppe kann durch Fluorid oder unter sauren Bedingungen gespalten werden. Eine 3-, 4-, 5- oder 6-substituierte 2-Nitrobenzyloxy- oder 2-, 3-, 5- oder 6-substituierte 4-Nitrobenzyloxy-Verknüpfungsgruppe kann durch eine Photonenquelle (Photolyse) gespalten werden. Eine 3-, 4-, 5- oder 6-substituierte 2-Alkoxyphenoxy- oder 2-, 3-, 5- oder 6-substituierte 4-Alkoxyphenoxy-Verknüpfungsgruppe kann durch  $\text{Ce}(\text{NH}_4)_2(\text{NO}_3)_6$  (Oxidation) gespalten werden. Ein  $\text{NCO}_2$  (Urethan)-Linker kann durch Hydroxid (Base), Säure oder  $\text{LiAlH}_4$  (Reduktion) gespalten werden. Eine 3-Pentenyl-, 2-Butenyl- oder 1-Butenyl-Verknüpfungsgruppe kann durch  $\text{O}_3$ ,  $\text{OsO}_4/\text{IO}_4^-$  oder  $\text{KMnO}_4$  (Oxidation) gespalten werden. Eine 2-[3-, 4- oder 5-substituierte Furyl]oxy-Verknüpfungsgruppe kann durch  $\text{O}_2$ ,  $\text{Br}_2$ , MeOH oder Säure gespalten werden.

[0181] Bedingungen für die Spaltung weiterer labiler Verknüpfungsgruppen schließen ein: tert.-Alkyloxy-Verknüpfungsgruppen können durch Säure gespalten werden; Methyl(dialkyl)methoxy- oder 4-substituierte 2-Alkyl-1,3-dioxan-2-yl-Verknüpfungsgruppen können durch  $\text{H}_3\text{O}^+$  gespalten werden; 2-Silylethoxy-Verknüpfungsgruppen können durch Fluorid oder durch Säure gespalten werden; 2-(X)-Ethoxy (wobei X = Keto, Esteramid, Cyano,  $\text{NO}_2$ , Sulfid, Sulfoxid, Sulfon ist)-Verknüpfungsgruppen können unter alkalischen Bedingungen gespalten werden; 2-, 3-, 4-, 5- oder 6-substituierte Benzyloxy-Verknüpfungsgruppen können durch Säure oder unter reduktiven Bedingungen gespalten werden; 2-Butenyl-Verknüpfungsgruppen können durch  $(\text{Ph}_3\text{P})_3\text{RhCl}(\text{H})$  gespalten werden; 3-, 4-, 5- oder 6-substituierte 2-Bromphenoxy-Verknüpfungsgruppen können durch Li, Mg oder BuLi gespalten werden; Methylthiomethoxy-Verknüpfungsgruppen können durch  $\text{Hg}^{2+}$  gespalten werden; 2-(X)-Ethoxy (wobei X = ein Halogen)-Verknüpfungsgruppen können durch Zn oder Mg gespalten werden; 2-Hydroxyethoxy-Verknüpfungsgruppen können durch Oxidation (z. B. mit  $\text{Pb}(\text{OAc})_4$ ) gespalten werden.

[0182] Bevorzugte Linker sind diejenigen, die durch Säure oder Photolyse gespalten werden. Mehrere der säurelabilen Linker, die für Festphasen-Peptidsynthese entwickelt worden sind, sind zur Verknüpfung von Markierungen mit MOIs nützlich. Einige dieser Linker sind bei einem kürzlich erschienenen Überblick von Lloyd-Williams et al. beschrieben (Tetrahedron 49: 11065–11133, 1993). Ein nützlicher Linker-Typ beruht auf p-Alkoxybenzylalkoholen, von denen zwei, 4-Hydroxymethylphenoxyessigsäure und 4-(4-Hydroxymethyl-3-methoxyphenoxy)buttersäure, von Advanced ChemTech (Louisville, KY) kommerziell erhältlich sind. Beide Linker können über eine Esterbindung mit dem Benzylalkohol an eine Markierung gebunden werden und über eine Amidbindung mit der Carbonsäure an ein aminhaltiges MOI. Markierungen, die durch diese Moleküle gebunden sind, werden vom MOI mit variierenden Konzentrationen in Trifluoressigsäure freigesetzt. Die Spaltung dieser Linker führt zur Freisetzung einer Carbonsäure auf der Markierung. Säurespaltung von Markierungen, die durch verwandte Linker gebunden sind, wie etwa 2-Dimethoxy-4'-(carboxymethyloxy)-benzhydrylamin (erhältlich von Advanced ChemTech in Fmoc-geschützter Form), führt zur Freisetzung eines Carbonsäureamids auf der freigesetzten Markierung.

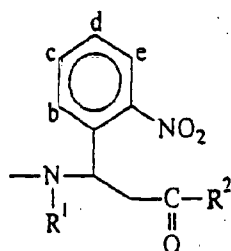
[0183] Die photolabilen Linker, die für diese Anwendung nützlich sind, sind ebenfalls größtenteils für die Festphasen-Peptidsynthese entwickelt worden (siehe Übersicht von Lloyd-Williams). Diese Linker beruhen üblicherweise auf 2-Nitrobenzylestern oder 2-Nitrobenzylamiden. Zwei Beispiele für photolabile Linker, über die vor kurzem in der Literatur berichtet worden ist, sind 4-(4-(1-Fmoc-Amino)ethyl)-2-methoxy-5-nitrophenoxy)butansäure (Holmes und Jones, J. Org. Chem. 60: 2318–2319, 1995) und 3-(Fmoc-Amino)-3-(2-nitrophenyl)propionsäure (Brown et al., Molecular Diversity 1: 4–12, 1995). Beide Linker können über die Carbonsäure an ein Amin auf dem MOI gebunden werden. Die Bindung der Markierung an den Linker erfolgt durch Bildung eines Amids zwischen einer Carbonsäure auf der Markierung und dem Amin auf dem Linker. Die Spaltung photolabiler Linker erfolgt üblicherweise mit UV-Licht mit 350 nm Wellenlänge bei Intensitäten und Zeiträumen, die den Fachleuten bekannt sind. Die Spaltung der Linker führt zu einer Freisetzung eines primären Amids auf der Markierung. Beispiele für photospaltbare Linker schließen Nitrophenylglycinester, exo- und endo-2-Benzonorborenylchloride und Methansulfonate und 3-Amino-3-(2-nitrophenyl)propionsäure ein. Beispiele für enzymatische Spaltung schließen Esterasen, die Esterbindungen spalten werden, Nuklearen, die Phosphodiester spalten werden, Proteasen, die Peptidbindungen spalten werden, etc. ein.

[0184] Eine bevorzugte Linkerkomponente besitzt eine ortho-Nitrobenzyl-Struktur, wie unten dargestellt:



in der ein Kohlenstoffatom an den Positionen a, b, c, d oder e mit  $-L^3-X$  substituiert ist und  $L^1$  (das vorzugsweise eine direkte Bindung ist) links von  $N(R^1)$  in der obigen Struktur vorhanden ist. Solch eine Linkerkomponente ist empfindlich gegenüber selektiver photo-induzierter Spaltung der Bindung zwischen dem Kohlenstoff mit der Bezeichnung "a" und  $N(R^1)$ . Die Identität von  $R^1$  ist typischerweise für die Spaltungsreaktion nicht kritisch,  $R^1$  wird jedoch vorzugsweise ausgewählt aus Wasserstoff und Hydrocarbyl. Die vorliegende Erfindung sorgt dafür, daß in der obigen Struktur  $-N(R^1)-$  durch  $-O-$  ersetzt werden könnte. Auch können in der obigen Struktur eine oder mehrere der Positionen b, e, d oder e fakultativ mit Alkyl, Alkoxy, Fluorid, Chlorid, Hydroxyl, Carboxylat oder Amid substituiert sein, wobei diese Substituenten bei jedem Auftreten unabhängig ausgewählt werden.

[0185] Eine weitere bevorzugte Linkerkomponente mit einem chemischen Griff  $L_n$  hat die folgende Struktur:



in der eine oder mehrere der Positionen b, c, d oder e mit Wasserstoff, Alkyl, Alkoxy, Fluorid, Chlorid, Hydroxyl, Carboxylat oder Amid substituiert ist,  $R^1$  Wasserstoff oder Hydrocarbyl ist und  $R^2$  -OH oder eine Gruppe ist, die eine Carbonsäure zur Kopplung mit einer weiteren Einheit entweder schützt oder aktiviert. Fluorkohlenstoff- und Fluorkohlenwasserstoffgruppen sind bevorzugte Gruppen, die Carbonsäure zu einer Kopplung mit einer anderen Einheit aktivieren.

### 3. Interessierendes Molekül (MOI)

[0186] Beispiele für MOIs schließen Nukleinsäuren oder Nukleinsäureanaloge (z. B. PNA), Fragmente von Nukleinsäuren (d. h. Nukleinsäurefragmente), synthetische Nukleinsäuren oder Fragmente, Oligonukleotide (z. B. DNA oder RNA), Proteine, Peptide, Antikörper oder Antikörperfragmente, Rezeptoren, Rezeptorliganden, Teile eines Ligandenpaares, Cytokine, Hormone, Oligosaccharide, synthetische organische Moleküle, pharmazeutische Wirkstoffe und Kombinationen derselben ein.

[0187] Bevorzugte MOIs schließen Nukleinsäurefragmente ein. Bevorzugte Nukleinsäurefragmente sind Primersequenzen, die zu den Vektoren vorhandenen Sequenzen komplementär sind, wobei die Vektoren zur Basensequenzierung verwendet werden. Vorzugsweise ist ein Nukleinsäurefragment direkt oder indirekt an eine Markierung an einem anderen als dem 3'-Ende des Fragments gebunden; am bevorzugtesten am 5'-Ende des Fragmentes. Nukleinsäurefragmente können gekauft oder hergestellt werden auf der Basis genetischer Datenbanken (z. B. Dib et al., Nature 380: 152–154, 1996 und CEPH Genotype Database, <http://www.cephb.fr>) und von kommerziellen Anbietern (z. B. Promega, Madison, WI).

[0188] Wie hierin verwendet, schließt MOI Derivate eines MOIs ein, die eine Funktioalität enthalten, die nützlich ist bei der Verknüpfung des MOIs mit einer T-L- $L_n$ -Verbindung. Ein Nukleinsäurefragment, das einen Phosphodiester am 5'-Ende aufweist, wobei der Phosphodiester auch an ein Alkylenamin gebunden ist, ist z. B. ein MOI. Solch ein MOI ist in z. B. U.S.-Patent 4,762,779 beschrieben, das hierin durch Bezugnahme mit einbezogen wird. Ein Nukleinsäurefragment mit einer internen Modifikation ist ebenfalls ein MOI. Eine beispielhafte interne Modifikation eines Nukleinsäurefragmentes ist, daß die Base (z. B. Adenin, Guanin, Cytosin, Thymidin, Uracil) modifiziert worden ist, um eine reaktive funktionelle Gruppe hinzuzufügen. Solche intern modifizierten Nukleinsäurefragmente sind kommerziell erhältlich von z. B. Glen Research, Herndon, VA. Eine weitere beispielhafte interne Modifikation eines Nukleinsäurefragmentes ist, daß ein abasisches Phosphoramidat verwendet wird, um einen modifizierten Phosphodiester zu synthetisieren, der zwischen einem Zucker und einer Phosphatgruppe eines Nukleinsäurefragmentes angeordnet ist. Das abasische Phosphoramidat enthält eine reaktive Gruppe, die ermöglicht, daß ein Nukleinsäurefragment, das diese Phosphoramidat-abgeleitete Einheit enthält, mit einer weiteren Einheit verknüpft wird z. B. mit einer T-L- $L_n$ -Verbindung. Solche abasischen Phosphorami-

date sind kommerziell erhältlich von z. B. Clonetech Laboratories, Inc., Palo Alto, CA.

#### 4. Chemische Griffe ( $L_n$ )

[0189] Ein chemischer Griff ist eine stabile, aber reaktive Atomanordnung, die als ein Teil eines ersten Moleküls vorhanden ist, wo der Griff eine chemische Reaktion mit einem komplementären chemischen Griff durchlaufen kann, der als Teil eines zweiten Moleküls vorhanden ist, um eine kovalente Bindung zwischen den zwei Molekülen zu bilden. Der chemische Griff kann z. B. eine Hydroxylgruppe sein und der komplementäre chemische Griff kann eine Carbonsäuregruppe (oder ein aktiviertes Derivat, z. B. ein Hydrofluoroarylester) sein, woraufhin eine Reaktion zwischen diesen zwei Griffen eine kovalente Bindung (spezifisch eine Estergruppe) bildet, die die zwei Moleküle miteinander verknüpft.

[0190] Chemische Griffe können in einer großen Anzahl von kovalente Bindungen bildenden Reaktionen verwendet werden, die für die Bindung von Markierungen an Linker und Linker in MOIs geeignet sind. Solche Reaktionen schließen Alkylierung (z. B. um Ether, Thioether zu bilden), Acylierung (z. B. um Ester, Amide, Carbamate, Harnstoffe, Thioharnstoffe zu bilden), Phosphorylierung (z. B. um Phosphate, Phosphonate, Phosphoramide, Phosphonamide zu bilden), Sulfonylierung (z. B. um Sulfonate, Sulfonamide zu bilden), Kondensation (z. B. um Imine, Oxime, Hydrazone zu bilden), Silylierung, Disulfidbildung und Erzeugung reaktiver Zwischenstufen, wie etwa Nitrene oder Carbene, durch Photolyse ein. Im allgemeinen sind Griffe und bindungsbildende Reaktionen, die zur Bindung von Markierungen an Linker geeignet sind, auch zur Bindung von Linkern an MOIs geeignet und umgekehrt. In einigen Fällen kann das MOI eine vorherige Modifikation oder Derivatisierung durchlaufen, um den für die Bindung des Linkers benötigten Griff bereitzustellen.

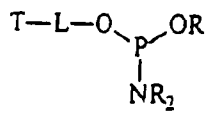
[0191] Ein Bindungstyp, der besonders nützlich ist zur Bindung von Linkern an MOIs, ist die Disulfidbindung. Ihre Bildung erfordert das Vorhandensein einer Thiolgruppe ("Griff") auf dem Linker und einer weiteren Thiolgruppe auf dem MOI. Milde oxidierende Bedingungen reichen dann aus, um die zwei Thiole als ein Disulfid miteinander zu verbinden. Disulfidbildung kann auch durch Verwendung eines Überschusses an einem geeigneten Disulfidaustauschreagens, z. B. Pyridyldisulfide, induziert werden. Weil Disulfidbildung leicht reversibel ist, kann das Disulfid auch als die spaltbare Bindung zur Freisetzung der Markierung verwendet werden, falls gewünscht. Dies wird typischerweise unter ähnlich milden Bedingungen durchgeführt, unter Verwendung eines Überschusses eines geeigneten Thiolaustauschreagens, z. B. Dithiothreitol.

[0192] Von besonderem Interesse für die Bindung von Markierungen (oder Markierungen mit Linkern) an Oligonukleotide ist die Bildung von Amidbindungen. Griffe aus primärem aliphatischen Amin können leicht mit Phosphoramiditen, wie etwa 6-Monomethoxytritylhexylcyanoethyl-N,N-düsoäpropylphosphoramidit (erhältlich von Glen Research, Sterling, VA), in synthetische Oligonukleotide eingeführt werden. Die in natürlichen Nukleotiden anzutreffenden Amine, wie etwa Adenosin und Guanosen, sind praktisch nicht-reaktiv, verglichen mit dem eingeführten primären Amin. Dieser Unterschied in der Reaktivität bildet die Grundlage für die Fähigkeit, selektiv Amide und verwandte Bindungsgruppen (z. B. Harnstoffe, Thioharnstoffe, Sulfonamide) mit dem eingeführten primären Amin und nicht mit den Nukleotid-Aminen zu bilden.

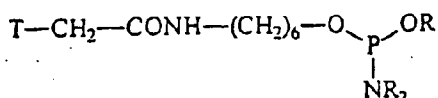
[0193] Wie in dem Katalog von Molecular Probes (Eugene, OR) aufgelistet, schließt eine teilweise Aufzählung von aminreaktiven funktionellen Gruppen aktivierte Carbonsäureester, Isocyanate, Isothiocyanate, Sulfonylhalogenide und Dichloätriazene ein. Aktivester sind hervorragende Reagentien zur Aminmodifizierung, da die gebildeten Amidprodukte sehr stabil sind. Diese Reagentien zeigen auch gute Reaktivität mit aliphatischen Aminen und geringe Reaktivität mit den Nukleotid-Aminen von Oligonukleotiden. Beispiele für Aktivester schließen N-Hydroxysuccinimidester, Pentafluorphenylester, Tetrafluorphenylester und p-Nitrophenylester ein. Aktivester sind nützlich, weil sie aus praktisch jedem Molekül hergestellt werden können, das eine Carbonsäure enthält. Verfahren zur Herstellung von Aktivestern sind aufgelistet in Bodansky (Principles of Peptide Chemistry (2. Aufl.), Springer-Verlag, London, 1993).

[0194] Die "X"-Gruppe in den durch  $T^{MS}-L-X$  bezeichneten Molekülen dient als ein chemischer Griff, der es ermöglicht, daß das Molekül an ein Biomolekül, z. B. ein Nukleinsäuremolekül angebracht wird. Die X-Gruppe ist ein Phosphoramidit oder eine H-Phosphonat-Gruppe. Wenn X eines dieser zwei Funktionalitäten ist, dann kann das T-L-X-Molekül an das Ende eines Oligonukleotids angefügt werden, das durch jedes der gut bekannten Phosphoramidit(auch als Phosphit-Triester bekannt) oder H-Phosphonat-Synthese-Methodologien für die Oligonukleotid-Synthese synthetisiert wurde, hinzugefügt werden.

[0195] Zum Beispiel, wenn X eine Phosphoramiditgruppe ist, kann das T-L-X-Molekül die Struktur die Formel aufweisen,



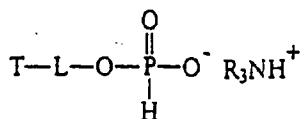
wobei ein bevorzugtes T-L-X-Molekül mit X als einem Phosphoramidit die Struktur



[0196] aufweist.

[0197] In den oben genannten Phosphoramidit-enthaltenden T-L-X-Molekülen ist R typischerweise eine Alkylgruppe, wie zum Beispiel eine C1-C6-Alkylgruppe, oder eine Alkylgruppe, die einen Substituenten anstelle eines Wasserstoff der Alkylgruppe aufweist, wobei geeignete einen Substituenten Cyano-(CN)-Gruppe einschließen. Daher kann "OR" im Phosphoramidit  $\text{OCH}_2\text{CH}_2\text{CN}$  sein und  $\text{NR}_2$  kann  $\text{N}(\text{isopropyl})_2$  sein, die die zwei am meisten bei der Herstellung von Oligonukleotiden unter der Verwendung von Phosphoramidit-Chemie angewendeten Gruppen sind.  $\text{NR}_2$  kann alternativ, zum Beispiel, eine Morpholingroup sein. In einer Ausführungsform ist R eine Alkylgruppe oder eine substituierte Alkylgruppe, die einen oder mehrere Substituenten ausgewählt aus Halogen und Cyano aufweist und die zwei R-Gruppen von  $\text{NR}_2$  können miteinander verbunden sein, um eine Cycloalkylgruppe zu bilden.

[0198] Ein zweiter Ansatz für die Synthese von Oligonukleotiden basiert vorteilhafterweise auf H-Phosphonat-Chemie. Ein T-L-X-Molekül der vorliegenden Erfindung kann einen chemischen Griff/X-Gruppe aufweisen, die eine H-Phosphonat-Gruppe ist und daher eine Struktur wie folgt aufweisen:



[0199] In der oben genannten H-Phosphonatgruppe stellt  $\text{R}_3$  drei Alkylgruppen dar, die typischerweise 1-6 Kohlenstoffatome in jeder Alkylgruppe aufweisen. Ethyl ist eine herkömmliche R-Gruppe in H-Phosphonat-Reagenzien, die in der Oligonukleotid-Synthese verwendet werden und ist daher eine bevorzugte R-Gruppe in T-L-X-Molekülen der Erfindung, worin X eine H-Phosphonat-Gruppe ist.

## 5. Linker-Bindung

[0200] Typischerweise wird ein einziger Linker-Typ verwendet, um einen bestimmten Satz oder eine bestimmte Familie von Markierungen mit einem bestimmten Satz oder einer bestimmten Familie von MOIs zu verbinden. In einer bevorzugten Ausführungsform der Erfindung kann ein einziger, gleichförmiger Prozeß befolgt werden, um alle verschiedenen T-L-MOI-Strukturen zu schaffen. Dies ist insbesondere dann vorteilhaft, wenn der Satz von T-L-MOI-Strukturen groß ist, weil es ermöglicht, daß der Satz unter Verwendung von Verfahren der kombinatorischen Chemie oder einer anderen Parallelverarbeitungstechnologie hergestellt werden kann. In einer ähnlichen Art und Weise ermöglicht die Verwendung eines einzigen Linker-Typs, das ein einziger, gleichförmiger Prozeß zur Spaltung aller verschiedenen T-L-MOI-Strukturen eingesetzt werden kann. Wieder ist dies für einen großen Satz von T-L-MOI-Strukturen vorteilhaft, weil der Satz auf eine parallele, wiederholte und/oder automatisierte Art und Weise verarbeitet werden kann.

[0201] Es gibt jedoch eine weitere Ausführungsform der vorliegenden Erfindung, in der zwei oder mehr Linker-Typen verwendet werden, um unterschiedliche Untersätze von Markierungen mit entsprechenden Untersätzen von MOIs zu verbinden. In diesem Fall können selektive Spaltungsbedingungen verwendet werden, um jeden der Linker unabhängig zu spalten, ohne die auf den anderen Untersätzen von MOIs vorhandenen Linker zu spalten.

[0202] Eine große Anzahl von kovalente Bindungen bildenden Reaktionen sind für die Bindung von Markierungen an Linker und Linker an MOIs geeignet. Solche Reaktionen schließen Alkylierung (z. B. um Ether, Thioether zu bilden), Acylierung (z. B. um Ester, Amide, Carbamate, Harnstoffe, Thioharnstoffe zu bilden), Phosphorylierung (z. B. um Phosphate, Phosphonate, Phosphoramide, Phosphonamide zu bilden), Sulfonylierung (z. B. um Sulfonate, Sulfonamide zu bilden), Kondensation (z. B. um Imine, Oxime, Hydrazone zu bilden), Silylierung, Disulfidbildung und Erzeugung reaktiver Zwischenstufen, wie etwa Nitrene oder Carbene, durch Photolyse ein. Im allgemeinen sind Griffe und bindungsbildende Reaktionen, die zur Bindung von Markierungen an Linker geeignet sind, auch zur Bindung von Linkern an MOIs geeignet und umgekehrt. In einigen Fällen kann das MOI eine vorherige Modifikation oder Derivatisierung durchlaufen, um den zur Bindung des Linkers notwendigen Griff bereitzustellen.

[0203] Ein Bindungstyp, der besonders nützlich für die Bindung von Linkern an MOIs ist, ist die Disulfidbindung. Ihre Bildung erfordert das Vorhandensein einer Thiolgruppe ("Griff") auf dem Linker und einer weiteren Thiolgruppe auf dem MOI. Milde oxidierende Bedingungen reichen dann aus, um die zwei Thiole miteinander als ein Disulfid zu verbinden. Disulfidbildung kann auch durch die Verwendung eines Überschusses eines geeigneten Disulfidaustauschreagens, z. B. Pyridyldisulfide, induziert werden. Weil Disulfidbildung leicht reversibel

ist, kann das Disulfid auch als die spaltbare Bindung zur Freisetzung der Markierung verwendet werden, falls gewünscht. Dies wird typischerweise unter ähnlich milden Bedingungen durchgeführt, unter Verwendung eines Überschusses eines geeigneten Thiolaustauschreagenz, z. B. Dithiothreitol.

[0204] Von besonderem Interesse für die Bindung von Markierungen an Oligonukleotide ist die Bildung von Amidbindungen. Griffe aus primärem aliphatischen Amin können leicht mit Phosphoramiditen, wie etwa 6-Monomethoxytritylhexylcyanoethyl-N,N-disopropylphosphoramidit (erhältlich von Glen Research, Sterling, VA), in synthetische Oligonukleotide eingeführt werden. Die in natürlichen Nukleotiden anzutreffenden Amine, wie etwa Adenosin und Guanosin, sind praktisch nicht-reaktiv, verglichen mit dem eingeführten primären Amin. Dieser Unterschied in der Reaktivität bildet die Grundlage für die Fähigkeit, selektiv Amide und verwandte Bindungsgruppen (z. B. Harnstoffe, Thioharnstoffe, Sulfonamide) mit dem eingeführten primären Amin und nicht mit den Nukleotid-Aminen zu bilden.

[0205] Wie in dem Katalog von Molecular Probes (Eugene, OR) aufgelistet, schließt eine teilweise Aufzählung von aminreaktiven funktionellen Gruppen aktivierte Carbonsäureester, Isocyanate, Isothiocyanate, Sulfonylhalogenide und Dichlortriazine ein. Aktivester sind hervorragende Reagentien zur Aminmodifizierung, da die gebildeten Amidprodukte sehr stabil sind. Diese Reagentien zeigen auch gute Reaktivität mit aliphatischen Aminen und geringe Reaktivität mit den Nukleotid-Aminen von Oligonukleotiden. Beispiele für Aktivester schließen N-Hydroxysuccinimidester, Pentafluorphenylester, Tetrafluorphenylester und p-Nitrophenylester ein. Aktivester sind nützlich, weil sie aus praktisch jedem Molekül hergestellt werden können, das eine Carbonsäure enthält. Verfahren zur Herstellung von Aktivestern sind aufgelistet in Bodansky (Principles of Peptide Chemistry (2. Aufl.), Springer-Verlag, London, 1993).

[0206] Zahlreiche kommerzielle Vernetzungsreagentien existieren, die als Linker dienen können (siehe z. B. Pierce Cross-linkers, Pierce Chemical Co., Rockford, IL). Unter diesen sind homobifunktionelle -aminreaktive Vernetzungsreagentien, wie beispielhaft homobifunktionelle Imidoester und N-Hydroxysuccinimidyl(NHS)-Ester. Es gibt auch heterobifunktionelle Vernetzungsreagentien, die zwei oder mehr unterschiedliche reaktive Gruppen besitzen, die sequentielle Reaktionen ermöglichen. Imidoester reagieren schnell mit Aminen bei alkalischem pH. NHS-Ester ergeben stabile Produkte, wenn sie mit primären oder sekundären Aminen umgesetzt werden. Maleimide, Alkyl- und Arylhalogenide, Alpha-Haloacyle und Pyridyldisulfide sind thiolreaktiv. Maleimide sind spezifisch für Thiol(Sulfhydryl)-Gruppen im pH-Bereich von 6,5 bis 7,5 und können bei alkalischem pH aminreaktiv werden. Die Thioether-Bindung ist unter physiologischen Bedingungen stabil. Alpha-Haloacetyl-Vernetzungsreagentien enthalten die Iodacetylgruppe und sind gegenüber Sulfhydrylen reaktiv. Die Imidazole können mit der Iodacetyleinheit reagieren, aber die Reaktion ist sehr langsam. Pyridyldisulfide reagieren mit Thiolgruppen, um eine Disulfidbindung zu bilden. Carbodiimide koppeln Carboxyle an primäre Amine von Hydraziden, was zur Bildung einer Acyl-Hydrazin-Bindung führt. Die Arylazide sind Photoaffinitätsreagentien, die chemisch inert sind, bis sie UV- oder sichtbarem Licht ausgesetzt werden. Wenn solche Verbindungen bei 250–460 nm photolysiert werden, wird ein reaktives Arylnitren gebildet. Das reaktive Arylnitren ist relativ unspezifisch. Glyoxale sind reaktiv gegenüber dem GuanidinyI-Teil von Arginin.

[0207] In einer typischen Ausführungsform der vorliegenden Erfindung wird eine Markierung zunächst an einen Linker gebunden, anschließend die Kombination von Markierung und Linker an ein MOI gebunden, um die Struktur T-L-MOI zu schaffen. Alternativ wird dieselbe Struktur gebildet, indem zunächst ein Linker an ein MOI gebunden wird und anschließend die Kombination von Linker und MOI an eine Markierung gebunden wird. Ein Beispiel ist, wenn das MOI ein DNA-Primer oder Oligonukleotid ist. In diesem Falle wird die Markierung typischerweise zunächst an einen Linker gebunden, anschließend das T-L an einen DNA-Primer oder ein Nukleotid gebunden, was dann z. B. in einer Sequenzierungsreaktion verwendet wird.

[0208] Eine nützliche Form, in der eine Markierung reversibel an ein MOI (z. B. ein Oligonukleotid oder einen DNA-Sequenzierungsprimer) gebunden werden kann, ist durch einen chemisch labilen Linker. Eine bevorzugte Konstruktion für den Linker ermöglicht, daß der Linker gespalten werden kann, wenn er einer flüchtigen organischen Säure ausgesetzt wird, z. B. Trifluoressigsäure (TFA). TFA ist besonders kompatibel mit den meisten Verfahren zur MS-Ionisierung, einschließlich Elektrospray.

[0209] Wie im Detail unten beschrieben, liefert die Erfindung eine Methodik für die Genotypisierung. Eine Zusammensetzung, die beim Genotypisierungsverfahren nützlich ist, umfaßt eine Mehrzahl von Verbindungen der Formel:

$T^{ms}$ -L-MOI

in der  $T^{ms}$  eine durch Massenspektrometrie nachweisbare organische Gruppe ist, die Kohlenstoff, wenigstens eines von Wasserstoff und Fluorid und fakultative Atome, die ausgewählt sind aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Iod, umfaßt. In der Formel ist L eine organische Gruppe, die ermöglicht, daß eine  $T^{ms}$ -enthaltende Einheit vom Rest der Verbindung abgespalten werden kann, wobei die  $T^{ms}$ -enthaltende Einheit eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterworfen wird, und ausgewählt ist aus tertiärem Amin, quartärem Amin und or-

ganischer Säure. In der Formel ist MOI ein Nukleinsäurefragment, wobei L an das MOI an einer anderen Stelle als dem 3'-Ende des MOI konjugiert ist. In der Zusammensetzung haben wenigstens zwei Verbindungen dieselbe  $T^{ms}$ , aber die MOI-Gruppen dieser Moleküle besitzen nicht-identische Nukleotidlängen.

[0210] Eine weitere Zusammensetzung, die beim Genotypisierungsverfahren nützlich ist, umfaßt eine Mehrzahl von Verbindungen der Formel:

#### $T^{ms}$ -L-MOI

in der  $T^{ms}$  eine durch Massenspektrometrie nachweisbare organische Gruppe ist, die Kohlenstoff, wenigstens eines von Sauerstoff und Fluorid und fakultative Atome, die ausgewählt sind aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Iod, umfaßt. In der Formel ist L eine organische Gruppe, die es ermöglicht, daß eine  $T^{ms}$ -enthaltende Einheit vom Rest der Verbindung abgespalten werden kann, wobei die  $T^{ms}$ -enthaltende Einheit eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterworfen wird, und ausgewählt ist aus tertiärem Amin, quartärem Amin und organischer Säure. In der Formel ist MOI ein Nukleinsäurefragment, wobei L an das MOI an einer anderen Stelle als dem 3'-Ende des MOI konjugiert ist. In der Zusammensetzung haben wenigstens zwei Verbindungen dieselbe  $T^{ms}$ , aber diese Verbindungen haben nicht-identische Elutionszeiten bei Säulenchromatographie.

[0211] Eine weitere Zusammensetzung, die beim Genotypisierungsverfahren verwendet werden kann, umfaßt eine Mehrzahl von Verbindungen der Formel:

#### $T^{ms}$ -L-MOI

in der  $T^{ms}$  eine durch Massenspektrometrie nachweisbare organische Gruppe ist, die Kohlenstoff, wenigstens eines von Sauerstoff und Fluorid und fakultative Atome, die ausgewählt sind aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Iod, umfaßt. In der Formel ist L eine organische Gruppe, die es ermöglicht, daß eine  $T^{ms}$ -enthaltende Einheit vom Rest der Verbindung abgespalten werden kann, wobei die  $T^{ms}$ -enthaltende Einheit eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterworfen wird, und ausgewählt ist aus tertiärem Amin, quartärem Amin und organischer Säure. In der Formel ist MOI ein Nukleinsäurefragment, wobei L an das MOI an einer anderen Stelle als dem 3'-Ende des MOI konjugiert ist. In der Zusammensetzung haben keine zwei Verbindungen, die dieselbe MOI-Nukleotidlänge aufweisen, auch dasselbe  $T^{ms}$ .

[0212] In der obigen Zusammensetzung ist die Mehrzahl vorzugsweise größer als 2 und vorzugsweise größer als 4. Auch hat das Nukleinsäurefragment im MOI eine zu einem Abschnitt eines Vektors komplementäre Sequenz, wobei das Fragment zum Priming einer Polynukleotidsynthese in der Lage ist. Vorzugsweise unterscheiden sich die  $T^{ms}$ -Gruppen von Mitgliedern der Mehrzahl um wenigstens 2 amu und können sich um wenigstens 4 amu unterscheiden.

[0213] Die Erfindung stellt auch eine Zusammensetzung zur Verfügung, die eine Mehrzahl von Sätzen von Verbindungen umfaßt, wobei jeder Satz von Verbindungen die Formel besitzt:

#### $T^{ms}$ -L-MOI

in der  $T^{ms}$  eine durch Massenspektrometrie nachweisbare organische Gruppe ist, die Kohlenstoff wenigstens eines von Wasserstoff und Fluorid und fakultative Atome, die ausgewählt sind aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Iod, umfaßt. In der Formel ist L eine organische Gruppe, die es ermöglicht, daß eine  $T^{ms}$ -enthaltende Einheit vom Rest der Verbindung abgespalten werden kann, wobei die  $T^{ms}$ -enthaltende Einheit eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterworfen wird, und ausgewählt ist aus tertiärem Amin, quartärem Amin und organischer Säure. In der Formel ist MOI auch ein Nukleinsäurefragment, wobei L an das MOI an einer anderen Stelle als dem 3'-Ende des MOI konjugiert ist. In der Zusammensetzung haben Mitglieder innerhalb eines ersten Satzes von Verbindungen identische  $T^{ms}$ -Gruppen, haben jedoch nicht-identische MOI-Gruppen mit unterschiedlichen Anzahlen von Nukleotiden im MOI, und es gibt wenigstens zehn Mitglieder im ersten Satz, wobei zwischen den Sätzen die  $T^{ms}$ -Gruppen sich um wenigstens 2 amu unterscheiden. Die Mehrzahl ist vorzugsweise wenigstens 3 und bevorzugter wenigstens 5.

[0214] Die Erfindung stellt auch eine Zusammensetzung zur Verfügung, die eine Mehrzahl von Sätzen von Verbindungen umfaßt, wobei jeder Satz von Verbindungen die Formel hat

#### $T^{ms}$ -L-MOI

in der  $T^{ms}$  eine durch Massenspektrometrie nachweisbare organische Gruppe ist, die Kohlenstoff, wenigstens



eines von Wasserstoff und Fluorid und fakultative Atome, die ausgewählt sind aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Iod, umfaßt. In der Formel ist L eine organische Gruppe, die es ermöglicht, daß eine T<sup>ms</sup>-enthaltende Einheit vom Rest der Verbindung abgespalten werden kann, wobei die T<sup>ms</sup>-enthaltende Einheit eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterworfen wird, und ausgewählt ist aus tertiärem Amin, quartärem Amin und organischer Säure. In der Formel ist MOI ein Nukleinsäurefragment, wobei L an das MOI einer anderen Stelle als dem 3'-Ende des MOI konjugiert ist. In der Zusammensetzung haben die Verbindungen innerhalb eines Satzes dieselbe Elutionszeit, aber nicht-identische T<sup>ms</sup>-Gruppen.

[0215] Zusätzlich stellt die Erfindung einen Kit zum Genotypisieren zur Verfügung. Der Kit umfaßt eine Mehrzahl von Amplifikationsprimerpaaren, wobei wenigstens einer der Primer die Formel hat:

T<sup>ms</sup>-L-MOI

in der T<sup>ms</sup> eine durch Massenspektrometrie nachweisbare organische Gruppe ist, die Kohlenstoff, wenigstens eines von Wasserstoff und Fluorid und fakultative Atome, die ausgewählt sind aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Iod, umfaßt. In der Formel ist L eine organische Gruppe, die es ermöglicht, daß eine T<sup>ms</sup>-enthaltende Einheit vom Rest der Verbindung abgespalten werden kann, wobei die T<sup>ms</sup>-enthaltende Einheit eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterworfen wird, und ausgewählt ist aus tertiärem Amin, quartärem Amin und organischer Säure. In der Formel ist MOI ein Nukleinsäurefragment, wobei L an das MOI an einer anderen Stelle als dem 3'-Ende des MOI konjugiert ist; und jedes Primerpaar assoziiert mit einem unterschiedlichen Locus. Im Kit ist die Mehrzahl vorzugsweise wenigstens 3 und bevorzugter wenigstens 5.

[0216] Wie oben angegeben, stellt die vorliegende Erfindung Zusammensetzungen und Verfahren zur Bestimmung der Sequenz von Nukleinsäuremolekülen zur Verfügung. Kurz gesagt umfassen solche Verfahren allgemein die Schritte: (a) Erzeugen markierter Nukleinsäurefragmente, die komplementär zu einem ausgewählten Nukleinsäuremolekül (z. B. markierte Fragmente) von einem ersten Ende bis zu einem zweiten Ende eines Nukleinsäuremoleküls sind, wobei eine Markierung mit einem bestimmten oder ausgewählten Nukleotid korrelativ ist und mit irgendeinem aus einer Vielzahl von Verfahren nachgewiesen werden kann, (b) Abtrennen der markierten Fragmente nach Sequenzlänge, (c) Abspalten einer Markierung von einem markierten Fragment und (d) Nachweisen der Markierungen und Bestimmen der Sequenz des Nukleinsäuremoleküls daraus. Jeder der Aspekte wird unten detaillierter diskutiert werden.

## B. DIAGNOSEVERFAHREN

### 1. Einführung

[0217] Wie oben angegeben, stellt die vorliegende Erfindung auch eine breite Vielfalt von Verfahren zur Verfügung, in denen die oben beschriebenen Markierungen und/oder Linker anstelle herkömmlicher Markierungen (z. B. radioaktiv oder enzymatisch) eingesetzt werden können, um die Spezifität, Empfindlichkeit oder Anzahl von Proben, die gleichzeitig analysiert werden können, in einem vorgegebenen Verfahren zu erhöhen. Repräsentative Beispiele für solche Verfahren, die verbessert werden können, schließen z. B. RNA-Amplifikation (siehe Lizardi et al., *Bio/Technology* 6: 1197–1202, 1988; Kramer et al., *Nature* 339: 401–402, 1989; Lomeli et al., *Clinical Chem.* 35(9): 1826–1831, 1989; U.S.-Patent Nr. 4,786,600) und DNA-Amplifikation unter Verwendung von LCR oder Polymerasekettenreaktion ("PCR") (siehe U.S. Patente Nm. 4,683,195, 4,683,202 und 4,800,159) ein.

[0218] Die CMST-Technologieplattform kann in einer Zahl von Anwendungen verwendet werden, in denen Nukleinsäure-Messungen in einem großen Maßstab durchgeführt werden. Diese Technologieplattform kann mit oder ohne einer Auftrennungs- oder Größen-Methodologie verwendet werden. Ein Beispiel eines nicht-Größen Assays würde Single Nucleotide Polymorphism (SNP) Assays einschließen, in denen Oligonukleotide verwendet werden, um die Anwesenheit oder Abwesenheit einer Basenveränderung in einer Ziel-Nukleinsäure nachzuweisen. Alternativ kann ein HPLC- oder Auftrennungssystem an den Massenspektrometrie-Detektor (MSD) angehängt werden, in dem Nukleinsäurefragmente nach ihrer Größe sortiert werden können und daher wird ein Massenspektrometrie-Marker mit einer Retentionszeit kombiniert, um eine Sequenz zu identifizieren. Die HPLC von Nukleinsäuren (Huber et al., 1993, *Anal. Biochem* 212, 5.351; Huber et al., 1993, *Nuc. Acids Res.*, 21, 5.1061; Huber et al., 1993, *Biotechniques*, 16, 5.898) oder denaturierende HPLC (DHPLC) ist ein bei der Auftrennung von DNA-Duplices, die sich in ihrer Identität um eine oder mehrere Basenpaare unterscheiden, allgemein erfolgreiches Verfahren und ist für das Screening von Mutanten allgemein brauchbar. Ein allgemeines Verfahren wurde unter der Verwendung von 100 mM Triethylammoniumacetat als dem Ionen-paarenden Reagenz entwickelt, in dem Oligonukleotide über HPLC erfolgreich auf alkylisiertem nicht-porösen 2,3 µM Poly(styrol-divinylbenzol) aufgetrennt wurden (Oefner et al., 1994, *Anal. Biochem.*, 223, S.39).

Diese Technik erlaubt auch die Auftrennung von PCR-Produkten, die sich in ihrer Länge in nur 4 bis 8 Basenpaaren innerhalb eines Größenbereichs von 50 bis 200 Nukleotide unterscheiden.

[0219] DHPLC hat die Suche nach SNPs signifikant beschleunigt (Oefner und Underhill, *Am. J. Hum. Genetics*, 57, A266, 1995). Zahlreiche Anwendungen der DHPLC-Technik schließen die Identifizierung von Polymorphismen auf dem menschlichen Y-Chromosom ein, um evolutionäre Untersuchungen zu erleichtern (PNAS USA, 93, 196–200, 1996), und die schnelle Identifizierung von Erkrankungen die Mutationen auf dem Chromosom 19 verursachen, die Ataxia verursachen (*Cell*, 87, 543–552, 1996).

[0220] Während es schwierig war genetische Tests für multifaktorielle Erkrankungen zu entwickeln, werden mehr als 200 bekannte menschliche Erkrankungen durch einen Defekt in einem einzelnen Gen verursacht, oftmals eine Veränderung eines einzelnen Aminosäurerests (Olsen, *Biotechnology: An industry comes of age*, National Academic Press, 1986).

[0221] Sensitive Mutations-Nachweistechiken bieten außergewöhnliche Möglichkeiten für das Mutationscreening. Effiziente genetische Tests können auch ein Screening auf onkogene Mutationen in Zellen ermöglichen, die aus dem respiratorischen Trakt oder der Blase abgestoßen werden, in Zusammenhang mit Gesundheitsvorsorgeuntersuchungen (Sidransky et al., *Science* 252: 706, 1991). Auch, wenn ein unbekanntes Gen eine genetische Erkrankung verursacht, sind Verfahren brauchbar, um DNA-Sequenzvarianten zu überwachen, um die Vererbung der Erkrankung durch genetische Kopplungsanalyse zu untersuchen. Mehrere verschiedene Ansätze wurden verfolgt, jedoch ist keiner weder effizient noch billig genug um wirklich weit verbreitet angewendet zu werden (Cotton. *RGH*, (1997), *Mutation Detection*, Oxford University Press, New York). Mutationen, die ein einzelnes Nukleotid einschließen, können in einer Probe durch physikalische, chemische oder enzymatische Mittel identifiziert werden. Im allgemeinen, können Verfahren zum Mutationsnachweis in Scann-Techniken (Fearon, 1997, *Science*, 278, p 1043–1050) eingeteilt werden, die geeignet sind, um bisher unbekannte Mutationen zu modifizieren und Techniken, die dazu aufgebaut sind, um bekannten Sequenzvarianten nachzuweisen, zu unterscheiden oder zu quantifizieren (Hotzman et al., 1997, *Science*, 278, 602–605).

[0222] Verschiedenen Scan-Techniken zum Mutationsnachweis wurden für Heteroduplexe entwickelt, wo die Anwesenheit eines Fehlpaars ein abnormales Verhalten induziert, wenn die Duplex teilweise denaturiert wird. Dieses Phänomen wird in den denaturierenden und Temperaturgradientengelelektrophoreseverfahren (jeweils DGGE und TGGE) ausgenutzt. Duplexe, die in sogar einer einzelnen Nukleotidposition fehlgepaart sind, können teilweise denaturieren, was zu verzögerter Wanderung führt, wenn in einem ansteigenden denaturierenden Gradientengelelektrophoret (Orita, *Genomics* 5: 874, 1989; Keen, *Trends Genet.* 7: 5, 1991, Myers et al., *Nature* 313: 495, 1985; Abrams et al., *Genomics* 7: 463, 1990; henco et al., *Nucl. Acids Res.* 18: 6733, 1990). Obwohl Mutationen nachgewiesen werden können, sind keine Informationen im Hinblick auf den genauen Ort oder die Sequenz um die Mutation herum erhalten.

[0223] Fehlgepaarte Basen in einer Duplex sind ebenfalls gegenüber chemischer Modifikation empfindlich. Solche Modifikation kann die Stränge gegenüber der Spaltung an der Stelle der Fehlpaarung empfindlich machen oder verursachen, daß eine Polymerase in einer anschließenden Verlängerungsreaktion anhält. Die chemische Spaltungstechnik erlaubt die Identifizierung einer Mutation in vielen Sequenzen von bis zu 2 kb und stellt Informationen über den ungefähren Ort der fehlgepaarten Nukleotide zur Verfügung (Cotton et al., *PNAS USA* 85: 4397, 1988; Ganguly et al., *Nucl. Acids Res.* 18: 3933, 1991).

[0224] Eine alternative Strategie zum Nachweis einer Mutation in einem DNA-Strang ist durch Substituierung (während der Synthese) eines der normalen Nukleotide in einem modifizierten Nukleotid, das das Molekulargewicht oder andere physikalische Parameter des Produkts verändert. Ein Strang mit einer erhöhten oder verringerten Zahl diesen modifizierten Nukleotids relativ zu der Wildtyp-Sequenz zeigt eine veränderte elektrophoretische Mobilität (Naylor et al., *Lancet* 337: 635, 1991). Wiederum weist diese Technik die Anwesenheit einer Mutation nach, stellt jedoch nicht den Ort zur Verfügung.

[0225] Alle der oben genannten Techniken zeigen die Anwesenheit einer Mutation in einem begrenzten Segment von DNA an, und einige dieser ermöglichen die ungefähre Lokalisierung innerhalb des Segments. Jedoch ist eine Sequenzanalyse immer noch erforderlich, um die genaue Position der Basenveränderung zu lokalisieren.

[0226] Eine große Anzahl von anderen Techniken wurden entwickelt, um bekannte Sequenzvarianten oder Single Nucleotide Polymorphismen zu analysieren. Die Automatisierung und Wirtschaftlichkeit und sehr wichtige Überlegungen für diese Typen von Analysen, die zum Screening von Individuen und der allgemeinen Population angewendet werden können. Mutationen können über ihre destabilisierenden Effekte auf die Hybridisierung von kurzen Oligonukleotidsonden an eine Zielsequenz identifiziert werden (siehe Wetmur, *Crit. Rev. Biochem. Mol. Biol.*, 26: 227, 1991). Im allgemeinen, schließt diese Technik, Allel-spezifische Oligonukleotid-hybridisierung, die Amplifikation von Zielsequenzen ein, und die anschließende Hybridisierung mit kurzen Oligonukleotidsonden. Der Oligonukleotid-Ligations-Assay ist eine Erweiterung von PCR-basiertem Screening, der einen ELISAbasierten Assay verwendet (OLA, Nickerson et al., *Proc. Natl. Acad. Sci. USA* 87: 8923, 1990), um die PCR-Produkte nachzuweisen, die Zielsequenz enthalten. Daher werden sowohl Gelelektrophorese und Koloniehybridisierung beseitigt.

[0227] Wie oben angegeben stellt die CMST-Technologie ebenfalls eine breite Vielzahl von Verfahren zur Verfügung, wo die absplittbaren Marker und/oder Linker anstelle von herkömmlichen Markierungen (z. B. radioaktiv, Fluoreszenz oder enzymatisch) verwendet werden können, um die Spezifität, Sensitivität oder Zahl von Proben, die gleichzeitig analysiert werden können, innerhalb eines bestimmten Verfahrens zu erhöhen. Repräsentative Beispiele von solchen Verfahren, die verbessert werden können, schließen zum Beispiel Standard-Nukleinsäurehybridisierungsreaktionen (siehe Sambrook et al., *supra*), diagnostische Reaktionen wie zum Beispiel die Cycling-Probe-Technology (CPT) (siehe U.S.-Patente Nrn. 4,876,187 und 5,011,769) oder Oligonukleotid-Ligations-Assay (OLA) (Burket et al., *Science* 196: 180, 1987).

[0228] Die CMST-Technologie, kombiniert mit Hybridisierung kann für die Forensik verwendet werden. Die DNA-Analyse ermöglicht leicht die Ableitung von Verwandtschaft zwischen Individuen, wie es bei Vaterschaftstests erforderlich ist. Die genetische Analyse hat sich als sehr brauchbar bei der Knochenmarkstransplantation erwiesen, wo es erforderlich ist zwischen engen Verwandten Donor- und Empfängerzellen zu unterscheiden. Zwei Typen von Sonden sind nun für DNA-Fingerprinting und -Typisierung in Gebrauch. Polymorphe Minisatelliten-DNA-Sonden identifizieren multiple DNA-Sequenzen, die jeweils in variablen Formen in verschiedenen Individuen vorhanden sind, wodurch Muster erzeugt werden, die komplex und zwischen den Individuen hoch variabel sind. VNTR-Sonden identifizieren einzelne Sequenzen in einem Genom, jedoch können diese Sequenzen in bis zu 30 verschiedene Formen in der menschlichen Population vorhanden sein, wie durch die Größe der identifizierten Fragmente unterschieden (Bennett und Todd, 1996, *Ann. Rev. Genetics*, 30, S. 343–70).

[0229] Tumordiagnose und -Phasenbestimmung (Goodfellow und Wells, 1995, *J. Natl. Cancer Inst.*, 87, S. 1515–23) ist eine weitere Anwendung der CMST-Technologie-Plattform. Der Nachweis von Onkogenen und deren jeweiligen Polymorphismen ist ein wichtiges Feld der Nukleinsäurediagnose. Die zellulären Onkogene können durch spezifische Modifikationen, wie zum Beispiel Punktmutationen, aktiviert werden (wie bei dem c-K-ras-Onkogen in Blasenkarzinom und in colorektalen Tumoren), Promotor-Induktion, Gen-Amplifikation (wie in dem N-myc-Onkogen im Fall von Neuroblastom) oder der Neuordnung von Chromosomen (wie bei der Translokation des c-abl-Onkogens von Chromosom 9 zu Chromosom 22 im Fall von chronischer Myeloid-Leukämie). Die CMST-Technologie kann auch für die Transplantationsanalyse, Genomdiagnose (4% aller Neugeborenen werden mit genetischen Defekten geboren) angewendet werden. Von den 3.500 beschriebenen Erbkrankheiten, die durch die Modifikation von nur einem einzelnen Gen verursacht werden, sind die primären molekularen Effekte nur für ungefähr 400 bekannt. Die Verwendung von DNA-Sonden mit absplittbaren Markern kann dazu verwendet werden, um die Anwesenheit oder Abwesenheit von Mikroorganismen in jedem Typ von Probe oder Spezimen nachzuweisen.

[0230] Die CMST-Technologie-Plattform kann mit verschiedenen größenabhängigen Techniken gekoppelt werden. Kapillarelektrophorese (CE) in ihren verschiedenen Ausführungen (freie Lösung, Isotachophorese, isoelektrische Fokussierung, Polyacrylamidgel, mizellare Elektrokinetik "Chromatographie") entwickelt sich als ein Verfahren für schnelle hochauflösende Abtrennungen von sehr kleinen Probenvolumen von komplexen Gemischen. In Kombination mit der inhärenten Aktivität und Selektivität von MS, ist CE-MS eine potentiell leistungsfähige Technik für die Bioanalyse. In der hier beschriebenen neuen Anwendung könnte die Überlappung dieser zwei Verfahren zu überlegenen DNA-Sequenzierungsverfahren führen, die die augenblickliche Rate-Verfahren der Sequenzierung um mehrere Größenordnungen übersteigt.

[0231] Die Korrespondenz zwischen CE und Elektrosprayionisierung-(ESI)-Flußraten und die Tatsache, daß beide durch (und primär verwendet werden für) ionische Spezies in Lösung unterstützt werden, stellt die Base für eine extrem attraktive Kombination zur Verfügung. Die Kombination von sowohl Kapillaronenelektrophorese (CZE) und Kapillarisotachophorese mit Quadrupolmassenspektrometer, basierend auf ESI, wurden beschrieben (Olivares et al., *Anal. Chem.* 59: 1230, 1987; Smith et al., *Anal. Chem.* 60: 436, 1988; Loo et al., *Anal. Chem.* 179: 404, 1989; Edmonds et al., *J. Chroma.* 474:21, 1989; Loo et al., *J. Microcolumn Sep.* 1: 223, 1989; Lee et al., *J. Chromatog.* 458: 313, 1988; Smith et al., *J. Chromatog.* 480: 211, 1989; Grese et al., *J. Am. Chem. Soc.* 111: 2835, 1989).

[0232] Das wirksamste Abtrennverfahren für DNA-Fragmente ist Polyacrylamidgelelektrophorese (PAGE), im allgemeinen in einem Scheibengelformat. Jedoch ist die hauptsächliche Begrenzung der augenblicklichen Technologie die relativ lange Zeit, die erforderlich ist, um die Gelelektrophorese von DNA-Fragmenten durchzuführen, die in den Sequenzreaktionen produziert wurden. Eine Erhöhung um eine Größenordnung (10-fach) kann mit der Verwendung von Kapillarelektrophorese erreicht werden, die ultradünne Gele verwendet. In Polyacrylamidgelen werden DNA-Fragmente gesiebt und wandern als eine Funktion der Länge und dieser Ansatz wurde nun auf CE angewendet. Eine bemerkenswerte Plattenzahl pro Meter wurde nun mit kreuz-vernetztem Polyacrylamid ( $10^{-7}$  Platten pro Meter, Cohen et al., *Proc. Natl. Acad. Sci., USA* 85: 9660, 1988) erreicht. Solche CE-Säulen, wie bisher beschrieben, können für die DNA-Sequenzierung verwendet werden. Smith und andere (Smith et al., *Nuc. Acids. Res.* 18: 4417, 1990) haben die parallele Verwendung von vielen Kapillaren vorgeschlagen, um den Durchsatz zu erhöhen. Ebenfalls haben Mathies und Huang (Mathies und Hang, *Nature* 359: 167, 1992) eine Kapillarelektrophorese eingeführt, in der die Abtrennungen über eine parallele Anordnung von Kapillaren durchgeführt wird, und eine Hochdurchsatz-Sequenzierung gezeigt (Huang et al., *Anal. Chem.*

64: 967, 1992, Huang et al., Anal. Chem 64: 2149, 1992). Da es keinen Grund gibt, parallele Spesen laufen zu lassen, besteht kein Grund ein Scheibengel zu verwenden. Daher kann man ein Röhrengelformat für das elektrophoretische Abtrennverfahren anwenden. Grossuran (Grossuran et al., Genet. Anal. Tech. Appl. 9: 9, 1992) hat gezeigt, daß ein beträchtlicher Vorteil erhalten wird, wenn ein Röhrengelformat anstelle eines Scheibengelformats verwendet wird. Dies liegt an der größeren Fähigkeit, Joulewärme in einem Röhrenformat abzuleiten, verglichen mit einem Scheibengel, was zu schnelleren Laufzeiten führt (um 50%) und sehr viel höhere Auflösung von hoch molekular Gewichts-DNA-Fragmenten (größer als 1.000 Nukleotide). Lange Ablesbarkeit ist bei der genomischen Sequenzierung wichtig. Daher weist die Verwendung von abspaltbaren Mackern bei der Sequenzierung den weiteren Vorteil auf, daß es dem Anwender ermöglicht wird, das am meisten effiziente und sensitive DNA-Abtrennverfahren zu verwenden, daß auch die höchste Auflösung besitzt.

[0233] Das hinter der Verwendung von mikro-fabrizierten Vonichtungen stehende Konzept ist die Fähigkeit, die Informationsdichte bei der Elektrophorese zu erhöhen, durch Miniaturisierung der Spurendimensionen auf ungefähr 100 Mikrometer. Die elektronische Industrie verwendet routinemäßig Mikro-Fabrikation, um die Schaltkreise mit Merkmalen herzustellen von weniger als einem Mikron Größe. Die augenblickliche Dichte von Kapillaren-Anays ist auf den äußeren Durchmesser der Kapillarenröhre begrenzt. Die Mikro-Fabrikation von Kanälen stellt eine höherer Array-Dichte zur Verfügung. Die Mikro-Fabrikation erlaubt auch physikalische Anordnungen, die mit Glasfasern nicht möglich nicht und koppelt die Kanäle direkt zu anderen Vonichtungen auf einem Chip. Ein paar Vonichtungen wurden auf Microchips für Abtrenntechnologien konstruiert. Ein Gaschromatograph (Teny e al., IEEE Trans. Electron Devic, ED-26: 1880, 1979) und ein Flüssigkeitschromatograph (Manz et al., Sens. Actuators B1: 249, 1990) wurden auf Siliziumchips hergestellt, jedoch wurden diese Vonichtungen nicht weiter verwendet. Verschiedene Gruppen haben die Abtrennung von fluoreszenten Farbstoffen und Aminosäuren auf Mikro-hergestellten Vonichtungen berichtet (Manz et al., J. Chromatography 593: 253, 1992, Effenhauser et al., Anal. Chem. 65:2637, 1993). Vor kurzem haben Woolley und Mathies (Woolley und Mathies, Proc. Natl. Acad. Scii. 91: 11348, 1994) gezeigt, daß Photolithographie und chemisches Ätzen dazu verwendet kann, um eine große Zahl von Abtrennkanälen auf Glassubstraten herzustellen. Diese Kanäle werden mit Hedroxyethylzellulose-(HEC)-Abtrennmatrices gefüllt.

[0234] Hybotrope werden in Tests und anderen Verfahren vorteilhaft angewendet, wobei ein markiertes Oligonukleotid an ein komplementäres oder semi-komplementäres (d. h. fast, jedoch nicht genau dieselbe Sequenz, wie das markierte ODN) Nukleinsäurefragment hybridisiert wird. Hybotrope sind genauer beschrieben in zum Beispiel U.S.-Patentanmeldungen Nrn. 60/026,621 (angemeldet 24. September 1996); 08/719,132 (angemeldet 24. September 1996); 08/933,924 (angemeldet 23. September 1997); 09/002,051 (angemeldet 31. Dezember 1997) und PCT Internationale Publikationsnr. WO 98/13527, die alle in ihrer Gesamtheit hier aufgenommen werden.

[0235] Die Beobachtung, daß  $\Delta-T_m$  sich als eine Funktion der Konzentration von Hybotropen nicht verändert, weist wesentliche Brauchbarkeit zur Verwendung in DNA, RNA oder Nukleinsäureamplifikation auf, die auf der Primerverlängerung durch Polymerase basieren (z. B. Polymerasekettenreaktion, siehe U.S.-Patente Nrn. 4,683,195; 4,683,202 und 4,800,159, Cycling Probe Technology, NASBA), Ligation (LCR, Ligationskettenreaktion) und RNA-Amplifikation (siehe Lizardi et al., Bio/Technologie 6: 1197, 1988; Kramer et al., Nature 339: 401, 1989; Lomeli et al., Clin. Chem. 35: 1826, 1989; U.S.-Patent Nr. 3,786,600). Die Beobachtung, daß WT (Wildtyp) und MT (Mutanten) 30-mer Oligonukleotide (30 verbundene Nukleotide im Oligonukleotid) auf der Basis von thermischen Schmelzen in 0,5 M LiTCA unterschieden werden kann, erlaubt die Möglichkeit einer wesentlichen Verbesserung bei der Priming-Effizienz in der PCR. In seiner augenblicklichen Anordnung wird der PCR-Puffer eher für die Polymerase aktiviert, als für ein spezifisches Priming gemeint ist, mit der Einführung der Technik wurden Bedingungen entwickelt, die die Performance der Polymerase über die Performance der Spezifität von Priming mit Oligonukleotiden bevorzugen. Daher stellt der PCR-Puffer, wie augenblicklich käuflich erhältlich, keinen hohen Spiegel von Stringenz von Hybridisierung von PCR-Primern zur Verfügung, noch unterstützt er diese.

[0236] Käuflich erhältliche PCR-Puffer werden im Hinblick auf das Schmelzverhalten von 24-mer Oligonukleotiden in sowohl den Wildtyp-(WT) und Mutanten-(MT)-Formen untersucht. Alternativ wird das Priming in einer hybotropen Lösung durchgeführt und die Kettenverlängerung wird in einem getrennten Puffer durchgeführt, der die Polymerase unterstützt. Zum Beispiel könnte eine Festphasen-PCR verwendet werden, wobei die feste Phase durch zwei Lösungen hindurch bewegt wird. Ein Priming würde in einiger geeigneten Konzentrationen von LiTCA oder TMAATCA auftreten und dann würde die Polymerasekettenreaktion in einem verschiedenen PCR-Puffer stattfinden, der die Polymerase enthält. Es ist auch möglich, die ersten paar Runden in der Amplifikation in einer Hybotrop basierten Hybridisierungslösung durchzuführen und die restlichen Runden in einem normalen PCR-Puffer durchzuführen (im allgemeinen sind nur die ersten paar Runden für die Spezifität wichtig).

[0237] Die Verwendung von abasisch modifizierten Oligonukleotiden erhöht auch die Spezifität des Primens in der PCR (siehe, z. B. U.S.-Patentanmeldungen Nrn. 60/026,621 (angemeldet 24. September 1996); 08/719,132 (angemeldet 24. September 1996); 08/933,924 (angemeldet 23. September 1997); 09/002,051

(angemeldet 31. Dezember 1997); und PCT Internationale Publikation Nr. WO 98/13527). Eine in einem Oligonukleotid eingefügte abasische Substitution verringert den HCT um 2,5°C. Zwei Oligonukleotidsonden, die drei abasische Stellen pro 24-mer enthalten, weisen eine HCT-Abnahme um 8°C relativ zu der nicht substituierten Kontrolle auf. Diese Abnahme in der HCT erhöht den Spiegel der Spezifität von Primern in der PCR-Reaktion dramatisch. Dies liegt an der Reduktion von falschen oder Miß-Primern während der ersten 10 Zyklen der PCR. Gemeint ist, daß die Enthalpie der abasischen substituierten Oligonukleotide relativ zu dem nicht substituierten Primer ansteigt, wodurch sich die Spezifität des Primers erhöht. Der Primer ist bevorzugterweise 6 bis 36 Basen lang und enthält 1 bis 6 abasische Stellen. Die abasischen Stellen sind bevorzugterweise 4, 5, 6, 7 oder 8 Nukleotide von einander getrennt und können bis zu 12 zu 24 Nukleotide von einander getrennt sein. Die Substitutionen sind bevorzugterweise am 3'-Ende des Primers gehäuft, um die Spezifität der Primerverlängerung durch Nukleinsäurepolymerasen sicherzustellen.

[0238] Weiterhin verringert die Konzentration einer abasischen Stelle in einem PCR-Primer und die Verwendung einer hybotropen Salzlösung, die einen hohen Enthalpiewert für die Primerduplex unterstützt, signifikant die  $\Delta$ -HTC der Primer-Duplex. Wie oben diskutiert, wenn die  $\Delta$ -HTC abnimmt, dann steigt der Stringenzfaktor an und ein hochdiskriminatorisches Priming der Polymerasekettenreaktion kann stattfinden. Dies sind Bedingungen, die für Multiplex-PCRs erforderlich sind. Der Ausdruck Multiplex betrifft die Fähigkeit, mehr als ein Set von Primern in einer PCR-Reaktion zu verwenden und multiple Produkte zu erzeugen, oder die Fähigkeit mehr als eine Zielnukleinsäure pro Set von PCR-Primern zu verwenden. Die Verwendung des Hybotrops Tetramethylammoniumtrichloracetat ist von besonderer Brauchbarkeit, da die Abhängigkeit des G + C-Gehalts auf  $T_m$  (Stabilität) neutralisiert wird.

[0239] Die beschriebenen hybotropen Lösungen werden dazu verwendet, um die Spezifität des Primers in der PCR zu erhöhen. Es gibt verschiedene Optionen im Hinblick auf einen Mechanismus durch den die Spezifität des Priming Schritts verbessert werden kann. Die erste ist durch die Verwendung eines festen Trägers, an den einer der PCR-Primer (kovalent) angebracht ist. Der feste Träger kann viele Formen annehmen, wie zum Beispiel Perlen, Membranen, usw. Der Priming Schritt kann in dem Hybotropen stattfinden und dann kann der feste Träger gewaschen werden und in einer Lösung bewegt werden, die die Polymerasekettenreaktion unterstützt. Der feste Träger wird dann zurück in das Nesstrop für die Priming Reaktion bewegt und der Zyklus wiederholt. Der Kreislauf des festen Trägers zwischen zwei verschiedenen Lösungen muß nur bis zu einer begrenzten Zahl von Zeiten (1-15 Zyklen) erfolgen, wonach es dem herkömmlichen Amplifikationszyklus in einem standardisiertem PCR-Puffer erlaubt werden kann, fortzuschreiten. Alternativ werden die Zielnukleinsäuren von Interesse zwischen der Priming Lösung und der Polymeraseverlängerungslösung unter der Verwendung elektrischer Felder (d. h. Elektrophorese) hin und her bewegt.

[0240] Die Verwendung von Hybotropen und/oder abasischen oder anukleosidischen Oligonukleotidsonden kann dazu verwendet werden, die Spezifität und Effizienz von isothermalen Anwendungen von Polymerasen für die Amplifikation von Nukleinsäuresequenzen zu erhöhen. Die Anwendungen von isothermalen Bedingungen zur Verwendung von Nukleinsäurepolymerasen schließen Nukleinsäuresequenzierung, Genotypisierung, Mutationsnachweis, Oligonukleotid-Ligationsassays, Mutationsnachweis und ähnliches ein.

[0241] Ein anderes Verfahren, das dazu verwendet wird, die Spezifität in Hybridisierungsreaktionen zu verbessern, erzeugt Basenfehlpaarungen unter der Verwendung von Basenanaloga um jedes der A-, G-, C- oder T-Nukleotide zu ersetzen. Die Forschung hat gezeigt, daß einige Primer, die eine Basenpaarfehlpaarung enthalten, eine erhöhte Spezifität aufweisen, wenn die Fehlpaarung an präzisen Stellen plaziert wird (siehe Wenham et al., *Clinical Chemistry* 37: 241, 1991; Newton et al., *Nucleic Acids Research* 17: 2503, 1989; Ishikawa et al., *Human Immunology* 42: 315, 1995). Jedoch sind auch Unterschiede von so wenig wie 0,5°C in den Schmelztemperaturen gleichfalls zwischen perfekt passenden Hybriden verbreitet und durch das selbe Hybrid mit einer einzelnen Basenfehlpaarung eingefügt (siehe Tibanyenda et al., *European Journal of Biochemistry* 139: 19, 1984; Werntges et al., *Nucleic Acids Research* 14: 3773, 1986). Sogar noch bessere Spezifität wurde zwischen einem und zwei fehlgepaarten Duplices bemerkt, als diese zwischen einer perfekt gepaarten Duplex und der selben Duplex mit einer einzelnen Fehlpaarung beobachtet wurde (siehe Guo et al., *Nature Biotechnology* 15: 31, 1997). Guo et al. fanden eine  $\Delta T_m$  von 4°C zwischen null und einer Fehlpaarung und eine  $\Delta T_m$  von 13°C zwischen einer und zwei aneinander grenzenden Fehlpaarungen für eine 20-mer Duplex. Jedoch, auch mit zwei Fehlpaarungen gibt es oft immer noch wenig Destabilisierung der Duplex. Diese Unfähigkeit, konsistent zwischen Fehlpaarungen zu unterscheiden, trägt zur fehlenden Spezifität bei PCR bei.

[0242] Die Verwendung von mehr als einer Basenfehlpaarung pro Hybridisierung, wobei mindestens ein Nukleotidanalogon verwendet wurde, wurde untersucht (siehe Guo et al., *Nature Biotechnology* 15: 331, 1997). In diesem Fall besteht die analoge Verbindung aus einer 3-Nitropyrrole-Ersatz der Purin- oder Pyrimidin-Basen. 3-Nitropyrrol hat die Fähigkeit minimale Wasserstoffbrücken mit allen vier Basen auszubilden (siehe Nichols et al., *Nature* 369: 492, 1994; Bergstorm et al., *Journal of the American Chemical Society* 117: 1201, 1995). Durch Einfügung einer künstlichen Fehlpaarung treten große Unterschiede in den Duplexschmelztemperaturen auf, die von ungefähr 5°C bis 15°C reichen, wobei die größte Differenz auftritt, wenn die Fehlpaarung in der Mitte des 15-mer hybridisierenden Oligos lokalisiert ist. Signifikante Unterschiede in  $\Delta T_m$  treten auf, wenn ein künst-

liches Nukleotid in eine Duplex eingeführt wird, die bereits eine Basenfehlpaarung enthält, was eine zwei-Fehlpaarungsduplex erzeugt. Das Ausmaß der Destabilisierung hängt von dem Typ von Basenfehlpaarung ab (z. B. G/T) und der Trennung zwischen den beiden Fehlpaarungen. Bei experimenteller Untersuchung reichte das Basenanalogon-Nukleotid von 1 bis 7 Basen von der 3'-Seite der Basenfehlpaarung, die in der Mitte des 15-mers gehalten wurde. Die Unterschiede in  $\Delta T_m$  für die drei verschiedenen Basen-fehlgepaarten 15-Mere reichten von einer 2°C Stabilisierung (nur im Falle der C/T-Fehlpaarung und wenn die Fehlpaarungen aneinander angrenzen), bis zu einer 7°C weiteren Destabilisierung, wobei die maximale Destabilisierung konsistenterweise bei einer 3 oder 4 Basenfehlpaarungstrennung auftrat (siehe Guo et al., *Nature Biotechnology* 15: 331, 1997).

[0243] Wenn zwei künstliche Fehlpaarungen eingefügt werden, beeinflußt die Umgebung der künstlichen Basen das Ausmaß der Destabilisierung stark. Die zwei künstlichen Fehlpaarungen wurden auf der Mitte einer 21-Mer-Duplex zentriert, beginnend mit einem Abstand von 6 Basenpaaren. Die Destabilisierung, oder  $\Delta T_m$ , ist mindestens 12°C verglichen mit der perfekt gepaarten Duplex. Der größte Unterschied von über 20°C tritt auf, wenn die zwei künstlichen Fehlpaarungen 10 Basen von einander beabstandet sind. Dieser Unterschied entspricht einer Helixwindung und zeigt an, daß eine Art von Interaktion zwischen den zwei künstlichen Basen auftritt, die die Stabilität der Duplex verringert.

[0244] Experimentell, wenn verwendete PCR-Primer eine oder zwei künstliche Fehlpaarungen zwischen dem Primer und der DNA-Probe enthielten, ergab die PCR Ergebnisse, wie sie für einen perfekt passenden Primer erwartet würden (siehe Gua et al. *Nature Biotechnology* 15: 331, 1997). Jedoch, wenn der Primer sowohl eine echte und eine künstliche Fehlpaarung enthielt, versagte die PCR dabei, irgendein meßbares Ergebnis zu produzieren. Hingegen produzierten die PCR-Reaktionen mit perfekt passenden und echten Fehlpaarungen alle meßbare Mengen von PCR-Produkt. Die selbe Untersuchung fand ähnliche Ergebnisse, wenn Hybridisierungs sonden verwendet wurden: diejenigen mit perfekter Übereinstimmung, echten Fehlpaarungen und künstlichen Fehlpaarungen lagerten sich an, während die Sonden, die künstliche und echte Fehlpaarungen enthielten, dies nicht taten. Diese Untersuchungen zeigten, daß eine größere Spezifität erzeugt wird, wenn künstliche Basenfehlpaarungen in Hybridisierungsreaktionen eingefügt werden, als wenn natürlich auftretende Basenfehlpaarungen auftreten. Diese thermodynamisch weniger stabil als eine perfekt passende Hybridisierungsreaktion und daher ist weniger wahrscheinlich, daß sie ein falsch positives Ergebnis in einem Test oder PCR hervorrufen. Interessanterweise ist jedoch der oben angegebene Unterschied in der thermodynamischen Stabilität für Duplexes, die nur eine künstliche Fehlpaarung enthalten, nicht in der experimentellen Situation manifestiert.

[0245] Ein weiteres Mittel zur Hervorrufung der Hybridisierungsunterscheidung ist durch Unterschiede in der Stabilität zwischen Hybridisierungsduplexes, die Nicks und Lücken enthalten. In diesen Reaktion werden Duplexes aus in Tandem-gestapelten kurzen Oligomeren gebildet, die an einen längeren Strang hybridisieren, die sich entweder kontinuierlich oder nicht-kontinuierlich anlagern, wobei eine Lücke von ein paar Basen zurückgelassen wird. Hybridisierungen, die zu einem Nick führen werden "Stapel-Hybridisierung" unterzogen, wobei ein anderer DNA-Strang über die Nick-Stelle hinweg hybridisiert. Stapel-Hybridisierung tritt nicht auf, wenn Lücken in den nicht-kontinuierlichen Oligomeren vorhanden sind. Das Stapeln hat den Effekt von erhöhter Unterscheidung, wie durch abnehmende Dissoziationsraten und größere thermodynamische Stabilität nachgewiesen, anders als bei den nicht-kontinuierlichen Gegenspielern (siehe Lane et al., *Nucleic Acids Res.* 25: 611, 1997). Thermodynamische Messungen zeigen, daß die Unterschiede zwischen der Hybridisierungs-gestapelte Duplexes-Standard-freie Energieveränderung ( $\Delta G$ ) und den lückenhaften Duplexes 1,4 bis 2,4 kcal/Mol sind. Daher kann eine Unterscheidung bei der Hybridisierung durch die Verwendung von vielen kurzen Sonden ermöglicht werden.

[0246] Die meisten der augenblicklich verwendeten Basenmimetika sind das Ergebnis des Streben nach einer universellen Base. Viele verwenden Nitroazol-Basenanaloga und haben eine verringerte Unterscheidung bei der Basenpaarung gezeigt. Eine Serie von Nitroazol-Nukleobasenanaloga wurden in einem Versuch untersucht, zusätzliche Einsicht in die Signifikanz von elektronischer Struktur und heterozyklischer Größe bei der Basenpaarung für die Entwicklung von effektiveren universellen Basen zu erhalten (siehe Bergstrom et al., *Nucleic Acids Res.* 25: 1935, 1997). In dieser Arbeit wurden die thermodynamischen Eigenschaften der Desoxyribonukleoside von 3-Nitropyrrol, 4-Nitropyrazol, 4-Nitroimidazol und 5-Nitroindol gemessen. Zum Vergleich wurden thermodynamische Messungen auch mit den Desoxyribonukleosiden von Hypoxanthin und Pyrazol sowie einem abasischem Spacer, 1,2-Didesoxyribose durchgeführt. Vier Oligonukleotide wurden für jedes modifizierte Nukleosid synthetisiert, um Duplexes zu erhalten, in denen jede der vier natürlichen Basen gegenüber dem Basenmimetikum plaziert war. Alle der analysierten Basenmimetika erwiesen sich als viel weniger stabil als die natürlichen Basenpaarungen (A + T:  $T_m = 65,7^\circ\text{C}$ , C + G:  $T_m = 70,5^\circ\text{C}$ ), wobei die  $T_m$ s von 35–46°C für 5-Nitroindol bis 18–29°C für die anderen analysierten Nitroazolbasen reichten. Die einzige Ausnahme war 4-Nitroimidazol, gepaart mit dGTP, wobei die  $T_m$  40,9°C war. Bei der Analyse der freien Energie für die Duplexschmelzung wurde von dem 3-Nitropyrrol-Basenmimetikum gefunden, welches die geringste Unterscheidung aufweist, wenn mit jeder der vier natürlich auftretenden Basen gepaart, mit einer gesamten  $\Delta G$  0,4 von

kcal/Mol. Die nächste am wenigsten unterscheidende Base war 5-Nitroindol mit einem  $\Delta G$  von 0,8 kcal/Mol. Beide dieser Werte sind weniger als die  $\Delta G$  von 1,1 kcal/Mol, die zwischen den natürlichen Basenpaarungen von A + T und G + C gefunden wurden. 4-Nitropyrazol zeigen eine leichte Präferenz für Paarung mit A, mit einem  $\Delta G$  1 kcal/Mol, die stabiler als die C, G und T freien Energien war. Zuletzt zeigte 4-Nitroimidazol eine hohe Selektivität für Paarung mit G (wie durch ihren hohen Tm-Wert nachgewiesen wurde), aufgrund der Fähigkeit des Imidazol N3 mit dem Desoxyguanosin N1 eine Wasserstoffbrücke auszubilden. Es sollte jedoch bemerkt werden, daß die oben genannten Werte von den nächsten Basennachbarn des Mimetikums abhängig sind. Weitere Untersuchungen veränderten die nächsten Nachbarn und fanden, daß 3-Nitropyrrol und 5-Nitroindol relative nicht-diskriminatorische Basenpaarungspartner sind.

[0247] Von Interesse wurden die Enthalpie- und Entropie-Veränderungen als einander folgend gefunden (d. h. eine große Enthalpie-Veränderung korreliert mit einer großen Entropie-Veränderung), unabhängig vom verwendeten Basenmimetikum, was impliziert, daß die Korrelation zwischen  $\Delta S$  und  $\Delta H$  unabhängig von der Art der Assoziation der Basen ist. Was beobachtet wurde war, daß Enthalpie-Veränderungen in den nicht-Wasserstoffbindenden Basenmimetika gefunden wurden. Die geringen Werte für die Entropie-Veränderung reflektiert das größere Ausmaß an Bewegungsfreiheit, das für Basen möglich ist, die nicht in die Duplex durch Wasserstoffbrücken-Interaktionen eingeschlossen sind. Die kleinen Enthalpie-Veränderungen reflektieren Veränderungen in den Wasserstoffbrücken-Interaktionen als ein Ergebnis des Verlust von Wasserstoffbrücken-Interaktionen für die Base gegenüber des Basenmimetikums. Falls eine natürliche Base in der Helix gestapelt verblieb, ohne einen gegenüberliegenden Wasserstoffbrücken-bindenden Partner, dann hat sie die Wasserstoffbrücken-Interaktion mit Wasser verloren, ohne einen neuen Donor-Akzeptor-Partner wiederzuerhalten.

[0248] Eine ähnliche Untersuchung, schloß die Untersuchung von azyklischen Nukleosidanaloga mit Carboxamido- oder Nitro-substituierten heterozyklischen Basen ein (siehe Aerschot et al., *Nucleic Acids Res.* 23: 4363, 1995). Die Verwendung von azyklischen Nukleosiden stattet die Konstrukte mit genug Flexibilität aus, um ein gutes Basenstapeln zu ermöglichen, sowie es den Basenmimetika zu ermöglichen, eine Orientierung zu erreichen, um am besten mit der entsprechenden Base Basenpaarung auszubilden. Die untersuchten heterozyklischen Basen schlossen ein: 4,5-Imidazoldicarboxamid, 4-Nitroimidazol und 5-Nitroindazol. Diese Komplexe wurden gegen azyklisches Hypoxanthin, 1-(2-(Desoxy-(D-ribofuranosyl)-3-nitropyrrol, 5-Nitroindol, und 2-(Desoxyinosin gemessen. Alle neuen azyklischen Komplexe wiesen Spuren von 7–20°C weniger als diejenigen auf, die für die natürlichen Basen beobachtet wurden. 5-Nitroindol, und 2-(Desoxyinosin gemessen. Alle neuen azyklischen Komplexe wiesen Spuren von 7–20°C weniger als diejenigen auf, die für die natürlichen Basen beobachtet wurden. 5-Nitroindazol, wenn gegen jedes der vier natürlichen Basen gepaart, wie die kleinste Divergenz mit  $\Delta T_m$  nur 2,2°C auf, während das 4-Nitroimidazol einen Unterschied von 8,0°C aufwies, wobei dG signifikant außerhalb der Linie der anderen drei Basen lag, wie oben ähnlich beobachtet wurde. Von den Referenzverbindungen wies Desoxyinosin eine  $\Delta T_m$  von 5,6°C auf, 5-Nitroindol  $\Delta T_m$  war 1,0°C, 1-(2-(Desoxy-(D-ribofuranosyl)-3-nitropyrrol wies eine  $\Delta T_m$  von 5,1°C auf und die  $\Delta T_m$  von azyklischen Hypoxanthin war 4,8°C. Jedoch zeigten alle Basenmimetika ungefähr die selbe Destabilisierung ( $\Delta T_m$  45°C), wenn in einem Oligo plazierte, das fast ausschließlich aus Adenosinen bestand, mit der Ausnahme von 4-Nitroimidazol und azyklischem Desoxyinosin, die  $T_m$ s von jeweils 7,0°C und 8,9°C aufwiesen.

[0249] Aerschot und Mitarbeiter untersuchten auf dem Effekt der Einfügung von mehrfachen Basenmimetika in ein Oligo (siehe Aerschot et al., *Nucleic Acids Res.* 23: 4363, 1995). Insgesamt sanken die Schmelztemperaturen insgesamt ab, jedoch am deutlichsten bei der Einfügung von drei Basenmimetika. Die Nitroindole zeigten jedoch die geringste Menge von Temperaturunterschied.

[0250] Ein anderes Basenmimetikum, 1-(2-(Desoxy-(D-ribofuranosyl)imidazol-4-carboxiamid (Nukleosid 1), ahmt bevorzugterweise dA-, sowie dC-Nukleoside (siehe Johnson et al., *Nucleic Acids Res.* 25: 559, 1997). Die Fähigkeit für sowohl dA und dC zu substituieren ergibt sich aus der Rotation um die Carboxiamid-/Imidazolbindung sowie der Bindung zwischen dem Imidazol- und Furanose-Ring. Wenn das Imidazol anti gegenüber der Furanose ist und die Carboxiamidgruppe anti zu dem Imidazol ist, ist das einzelne Paar auf dem Sauerstoff und einem der Amid-NH-Wasserstoffe in einer Position, die das  $NH_2$  und N-1 von Adenosin nachahmt. Die Imidazolrotation um die glycosidische Bindung zu der Sinnorientierung plazierte die Amidgruppe in einer Position, die ungefähr den Positionen des  $NH_2$  und N-3 von Cytosin entspricht.

[0251] Wenn Nukleosid 1 für jedes natürlich auftretende Nukleosid substituiert wird, erhöht sich die Enthalpie, mit der größten Erhöhung für eine dG-Substitution für die 1-C-Paarung (von  $\Delta H = 74,7$  (kcal/Mol)/  $\Delta G = -16,5$  (kcal/Mol) für die G/C-Paarung von  $\Delta H = -45,5$  (kcal/Mol)/  $\Delta G = -5,8$  (kcal/Mol)). Die kleinste Enthalpie-Veränderung tritt für eine dA-Substitution auf ( $\Delta H -72,9$  (kcal/Mol)/  $\Delta G = -15,4$  (kcal/Mol) für A/T-Paarung von  $\Delta H = -66,7$  (kcal/Mol)/  $\Delta G = -11,7$  (kcal/Mol) für die 1-T-Paarung. Entsprechend nimmt die  $T_m$  signifikant von 65,7°C und 70,5°C für jeweils die A-T- und C-G-Paare ab auf 46,6°C für die 1-T-Paarung, 43,4°C für 1-G, 27,6°C für 1-A und 14,6°C für 1-C.

[0252] Wenn in einer PCR-Reaktion verwendet, werden Nukleosid 1 und sein N-Propylderivat bevorzugterweise als dATP-Analoga inkorporiert (siehe Sala et al., *Nucleic Acids Res.* 24: 3302, 1996). Jedoch, sobald in ein DNA-Templat inkorporiert, gab ihr mehrdeutiges Wasserstoffbindungspotential eine Miß-Inkorporation je-

der der natürlich auftretenden Basen bei Sequenzen von  $3 \times 10^{-2}$  pro Base pro Amplifikation. Die meisten der Substitutionen (die primär aus G bestanden) waren ein Ergebnis der Rotation um die Carboxyamidgruppe, wenn Teil des Templats. Zwischen 11 bis 15% der Substitutionen lagen an der Rotation der Imidazolgruppe um die glycosidische Bindung. Als Teil eines DNA-Templats wirkte das N-Propylderivat auf die selbe Weise wie 1, trotz seiner Propylgruppe. Diese Untersuchung zeigt, daß während 1 bevorzugterweise als dATP wirkt, es die Fähigkeit in einer PCR-Typ-Umgebung aufweist, sich auch als alle vier natürlich auftretenden Nukleotide zu benehmen. Aus dieser und den oben genannten Untersuchungen ist ersichtlich, daß ein breiter Bereich von Duplex-Stabilität durch Variationen in Basenmimetika und deren Plazierung innerhalb eines Oligonukleotids erhalten werden kann.

[0253] In einem Aspekt der vorliegenden Erfindung werden Verfahren zur Bestimmung der Identität eines Nukleinsäuremoleküls oder -fragments (oder zum Nachweisen des Vorhandenseins eines ausgewählten Nukleinsäuremoleküls oder -fragments) zur Verfügung gestellt, die die Schritte umfassen: (a) Erzeugen markierter Nukleinsäuremoleküle aus einem oder mehreren ausgewählten Target-Nukleinsäuremolekülen, wobei eine Markierung mit einem bestimmten Nukleinsäuremolekül korrelativ und mit Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie nachweisbar ist, (b) Abtrennen der markierten Moleküle nach Größe, (c) Abspalten der Markierungen von den markierten Molekülen und (d) Nachweisen der Markierungen mit Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie und Bestimmen der Identität der Nukleinsäuremoleküle daraus.

[0254] In einem verwandten Aspekt der Erfindung werden Verfahren zum Nachweisen eines ausgewählten Nukleinsäuremoleküls zur Verfügung gestellt, die die Schritte umfassen: (a) Kombinieren markierter Nukleinsäuresonden mit Target-Nukleinsäuremolekülen unter Bedingungen und für einen Zeitraum, die ausreichend sind, um Hybridisierung einer markierten Nukleinsäuresonde an eine komplementäre ausgewählte Target-Nukleinsäuresequenz zu ermöglichen, wobei eine markierte Nukleinsäuresonde mit Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie nachweisbar ist, (b) Verändern der Größe hybridisierter markierter Sonden, nicht-hybridisierter Sonden oder Target-Moleküle oder der Sonden:Target-Hybride, (c) Abtrennen der markierten Sonden nach Größe, (d) Abspalten der Markierungen von den markierten Sonden und (e) Nachweisen der Markierungen mit Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie und Bestimmen des ausgewählten Nukleinsäuremoleküls daraus. Diese und andere verwandte Techniken werden detaillierter unten diskutiert.

## 2. PCR

[0255] PCR kann eine gewünschte DNA-Sequenz eines beliebigen Ursprungs (Virus, Bakterium, Pflanze oder Mensch) Hunderte von Millionen mal innerhalb einiger Stunden amplifizieren. PCR ist insbesondere wertvoll, weil die Reaktion hoch-spezifisch, leicht zu automatisieren und zum Amplifizieren kleiner Probenmengen fähig ist. Aus diesen Gründen hat PCR eine große Bedeutung für die klinische Medizin, die Diagnose genetischer Erkrankungen, die forensische Wissenschaft und die Evolutionsbiologie gehabt.

[0256] Kurz gesagt ist PCR ein Verfahren, das auf einer spezialisierten Polymerase beruht, die einen komplementären Strang zu einem gegebenen DNA-Strang in einer Mischung synthetisieren kann, die die 4 DNA-Basen und 2 DNA-Fragmente (Primer, jeder etwa 20 Basen lang), die die Target-Sequenz flankieren, enthält. Die Mischung wird erhitzt, um die Stränge aus doppelsträngiger DNA, die die Target-Sequenz enthalten, voneinander, und anschließend abgekühlt, um zu ermöglichen, daß (1) die Primer ihre komplementären Sequenzen auf den voneinander getrennten Strängen finden und sich an sie binden, und (2) daß die Polymerase die Primer zu neuen komplementären Strängen verlängert. Wiederholte Zyklen von Erhitzen und Abkühlen multiplizieren die Target-DNA exponentiell, da jeder neue Doppelstrang sich so auftrennt, daß er zwei Matrizen für eine weitere Synthese wird. In etwa 1 Stunde können 20 PCR-Zyklen das Target um ein Millionenfaches amplifizieren.

[0257] In einer Ausführungsform der Erfindung werden Verfahren zum Bestimmen der Identität eines Nukleinsäuremoleküls oder zum Nachweisen des ausgewählten Nukleinsäuremoleküls in z. B. einer biologischen Probe unter Verwendung der PCR-Technik zur Verfügung gestellt. Kurz gesagt umfassen solche Verfahren die Schritte: Erzeugen einer Reihe markierter Nukleinsäurefragmente oder -moleküle während der PCR und Auftrennen der resultierenden Fragmente nach Größe. Der Größenauftrennschritt kann unter Verwendung irgendeiner der hierin beschriebenen Techniken durchgeführt werden, einschließlich z. B. Gelelektrophorese (z. B. Polyacrylamidgelelektrophorese) oder vorzugsweise HPLC. Die Markierungen werden anschließend von den aufgetrennten Fragmenten abgespalten und mit der entsprechenden Nachweisttechnologie nachgewiesen. Beispiele für solche Technologien sind hierin beschrieben worden und schließen z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV-Spektrometrie ein.

## 3. RNA-Fingerprinting und Differential Display

[0258] Wenn die Matrice RNA ist, ist der erste Schritt beim Fingerprinting reverse Transkription. Liang und Pardee (Science 257: 967, 1992) waren die ersten, die ein RNA-Fingerprinting-Protokoll beschrieben haben,



das einen Primer für reverse Transkription auf der Basis von Oligo(dT) verwendet, aber mit einem "Anker" aus zwei Basen am 5'-Ende (z. B. Oligo-5'-(dT<sub>11</sub>)CA-3'). Priming tritt hauptsächlich am 5'-Ende des Poly(rA)-Schwanzes auf und hauptsächlich in Sequenzen, die 5'-UpG-poly(rA)-3' enden, mit einer Selektivität, die sich einer aus 12 polyadenylierten RNAs nähert. Nach reverser Transkription und Denaturierung wird willkürliches Priming auf dem resultierenden ersten Strang aus cDNA durchgeführt. PCR kann jetzt verwendet werden, um einen Fingerprint von Produkten zu erzeugen, die am besten mit den Primern zusammenpassen und die vom 3'-Ende der mRNAs und polyadenylierten heterogenen RNAs abgeleitet werden. Dieses Protokoll ist als "Differential Display" bezeichnet worden.

[0259] Alternativ kann ein willkürlicher Primer im ersten Schritt der reversen Transkription verwendet werden, wobei diejenigen für die RNA-internen Regionen ausgewählt werden, die 6-8 Basenmatches mit dem 3'-Ende des Primers haben. Dem folgt willkürliches Priming des resultierenden ersten Stranges aus cDNA mit demselben oder einem unterschiedlichen willkürlichen Primer und anschließend PCR. Dieses besondere Protokoll sampled irgendwo in der RNA, einschließlich offene Leserahmen (Welsh et al., Nuc. Acids. Res. 20: 4965, 1992). Zusätzlich kann es auf RNAs verwendet werden, die nicht polyadenyliert sind, wie etwa viele bakterielle RNAs. Diese Variante von RNA-Fingerprinting durch PCR mit willkürlichem Primer ist RAP-PCR genannt worden.

[0260] Wenn Fingerprinting von RNA durch PCR mit willkürlichem Primer auf Proben durchgeführt wird, die aus Zellen, Geweben oder anderem biologischen Material gewonnen sind, die verschiedenen experimentellen Behandlungen unterzogen worden sind oder verschiedene Entwicklungsgeschichten haben, können Unterschiede in der Genexpression zwischen den Proben nachgewiesen werden. Für jede Reaktion wird angenommen, daß dieselbe Anzahl von effektiven PCR-Verdopplungsereignissen auftritt und alle Unterschiede in den anfänglichen Konzentrationen von cDNA-Produkten als ein Verhältnis von Intensitäten im letztendlichen Fingerprint konserviert werden. Es gibt keine sinnvollen Beziehungen zwischen den Intensitäten von Banden innerhalb einer einzigen Spur auf einem Gel, die eine Funktion von Match und Menge sind. Jedoch wird das Verhältnis zwischen Spuren für jede gesamplete RNA konserviert, was ermöglicht, daß differentiell exprimierte RNAs nachgewiesen werden können. Das Verhältnis von Ausgangsmaterialien zwischen Proben wird aufrechterhalten, selbst wenn die Anzahl von Zyklen ausreichend ist, um zu ermöglichen, daß die PCR-Reaktion absättigt. Dies beruht darauf, daß die Anzahl von Verdopplungen, die benötigt wird, um Sättigung zu erreichen, beinahe vollständig durch die invarianten Produkte gesteuert wird, die den Großteil des Fingerprint ausmachen. In dieser Hinsicht ist PCR-Fingerprinting verschieden von herkömmlicher PCR eines einzelnen Produkts, bei der das Verhältnis von Ausgangsmaterialien zwischen Proben nicht konserviert wird, sofern nicht Produkte in der exponentiellen Phase der Amplifikation gesampled werden.

[0261] In einer Ausführungsform der Erfindung werden Verfahren zum Bestimmen der Identität eines Nukleinsäuremoleküls oder zum Nachweisen eines ausgewählten Nukleinsäuremoleküls in z. B. einer biologischen Probe unter Verwendung der Technik des RNA-Fingerprinting zur Verfügung gestellt. Kurz gesagt umfassen solche Verfahren allgemein die Schritte: Erzeugen einer Reihe markierter Nukleinsäurefragmente. Die Fragmente, die mit PCR oder ähnlichen Amplifikationsschemen erzeugt werden, werden dann anschließend nach Größe getrennt. Der Schritt der Größentrennung kann z. B. jede der hierin beschriebenen Techniken sein, einschließlich z. B. Gelelektrophorese (z. B. Polyacrylamidgelelektrophorese) oder vorzugsweise HPLC. Die Markierungen werden dann von den aufgetrennten Fragmenten abgespalten, und anschließend werden die Markierungen mit der entsprechenden Nachweisttechnologie nachgewiesen. Repräsentative Beispiele für geeignete Technologien schließen Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV-Spektrometrie ein. Die relativen Mengen aller gegebenen Nukleinsäurefragmente sind nicht wichtig, aber die Größe der Bande ist informativ, wenn sie mit einer Kontrollprobe in Beziehung gesetzt wird.

#### 4. PCR-Einzelstrangkonnformationspolymorphismus auf Fluoreszenz-Basis (PCR-SSCP)

[0262] Eine Reihe von Verfahren, zusätzlich zum RFLP-Ansatz, sind für die Analyse von Basensubstitutionspolymorphismen verfügbar. Orita et al. haben einen Weg zur Analyse dieser Polymorphismen auf der Basis von Konformationsunterschieden denaturierter DNA entwickelt. Kurz gesagt wird Restriktionsenzymverdau oder PCR verwendet, um relativ kleine DNA-Fragmente zu erzeugen, die anschließend denaturiert und durch Elektrophorese auf nicht-denaturierenden Polyacrylamidgelen aufgetrennt werden. Konformationsunterschiede in den einzelsträngigen DNA-Fragmenten, die aus Basensubstitutionen resultieren, werden durch Elektrophoresemobilitätsverschiebung nachgewiesen. Basenpaarung innerhalb des Strangs schafft Einzelstrangkonnformationen, die hoch sequenzspezifisch und in der Elektrophoresemobilität unterschiedlich sind. Nachweisraten in verschiedenen Studien unter Verwendung von herkömmlichem SSCP reichen jedoch von 35% bis nahezu 100%, wobei die höchsten Nachweisraten am häufigsten mehrere unterschiedliche Bedingungen erfordern. Im Prinzip könnte das Verfahren auch verwendet werden, um Polymorphismen auf der Basis kurzer Insertionen oder Deletionen zu analysieren. Dieses Verfahren ist eines der leistungsfähigsten Tools zur Identifizierung von Punktmutationen und -deletionen in DNA (SSCP-PCR, Dean et al., Cell 61: 863, 1990).

[0263] In einer Ausführungsform der Erfindung werden Verfahren zum Bestimmen der Identität eines Nukleinsäuremoleküls oder zum Nachweisen eines ausgewählten Nukleinsäuremoleküls in z. B. einer biologischen Probe unter Verwendung der PCR-SSP-Technik zur Verfügung gestellt. Kurz gesagt umfassen solche Verfahren allgemein die Schritte: Erzeugen einer Reihe markierter Nukleinsäurefragmente. Die mit PCR erzeugten Fragmente werden dann nach Größe getrennt. Vorzugsweise ist der Größentrennschritt nicht-denaturierend, und die Nukleinsäurefragmente werden vor der Trennmethode denaturiert. Der Größentrennschritt kann z. B. mit Gelelektrophorese (z. B. Polyacrylamidgelelektrophorese) oder vorzugsweise HPLC durchgeführt werden. Markierungen werden anschließend von den abgetrennten Fragmenten abgespalten, und anschließend werden die Markierungen mit der entsprechenden Nachweistechnologie nachgewiesen (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV-Spektrometrie).

#### 5. Didesoxy-Fingerprinting (ddF)

[0264] Ein weiteres Verfahren ist beschrieben worden (ddF, Sarkar et al., Genomics 13: 441, 1992), das 100% von Einzelbasenveränderungen im Gen des Humanfaktors IX nachgewiesen hat, wenn getestet in einer retrospektiven und prospektiven Art und Weise. Insgesamt wurden 84 von 84 unterschiedlichen Sequenzveränderungen nachgewiesen, wenn genomische DNA von Patienten mit Haemophilie B analysiert wurde.

[0265] Kurz gesagt ist, bei den Anwendungen von Markierungen zum Genotypisieren oder anderen Zwecken, ein Verfahren, das verwendet werden kann, Didesoxy-Fingerprinting. Dieses Verfahren setzt einen Didesoxy-terminator in einer Sanger-Sequenzierungsreaktion ein. Das Prinzip des Verfahrens ist wie folgt: Eine Target-Nukleinsäure, die sequenziert werden soll, wird in eine Reaktion gegeben, die einen Didesoxy-Terminator aufweist, der komplementär zu der Base ist, von der bekannt ist, daß sie in der Target-Nukleinsäure mutiert ist. Wenn die Mutation z. B. zu einer A → G-Änderung führt, würde die Reaktion in einer C-Didesoxy-Terminator-Reaktion durchgeführt werden. PCR-Primer werden verwendet, um die interessierende Targetsequenz zu lokalisieren und zu amplifizieren. Wenn die hypothetische Targetsequenz die A → G-Änderung enthält, wird die Größe einer Population von Sequenzen aufgrund des Einbaus eines Didesoxy-Terminators in die amplifizierten Sequenzen verändert. In dieser besonderen Anwendung von Markierungen würde ein Fragment erzeugt werden, das eine vorhersagbare Größe im Fall einer Mutation besäße. Die Markierungen wären an das 5'-Ende der PCR-Primer gebunden und würden eine "Karte" für den Probenotyp und den Didesoxy-Terminortyp liefern. Eine PCR-Amplifikationsreaktion würde stattfinden, die resultierenden Fragmente würden nach Größe mit z. B. HPLC oder PAGE getrennt werden. Am Ende des Trennverfahrens werden die DNA-Fragmente in einem temporären Bezugsrahmen gesammelt, die Markierungen werden abgespalten, und das Vorhandensein oder Nichtvorhandensein von Mutation wird durch die Kettenlänge aufgrund frühzeitigen Kettenterminators durch den Einbau eines gegebenen Didesoxy-Terminators bestimmt.

[0266] Es ist wichtig anzumerken, daß ddF zu einem Gewinn oder Verlust eines Didesoxy-Terminationssegments und/oder einer Verschiebung der Mobilität wenigstens eines der Terminationssegmente oder -produkte führt. Daher wird bei diesem Verfahren eine Suche nach der Verschiebung von einer Fragmentmobilität vor einem hohen Hintergrund an Fragmenten mit anderer relativer Molekülmasse durchgeführt. Ein Vorteil ist die Vorkenntnis der Länge des Fragmentes, die mit einer gegebenen Mutation assoziiert ist.

[0267] In einer Ausführungsform der Erfindung werden Verfahren zum Bestimmen der Identität eines Nukleinsäuremoleküls oder zum Nachweisen eines selektierenden Nukleinsäuremoleküls in z. B. einer biologischen Probe unter Verwendung der ddF-Technik zur Verfügung gestellt. Kurz gesagt umfassen solche Verfahren allgemein die Schritte: Erzeugung einer Reihe markierter Nukleinsäurefragmente, gefolgt von der Trennung auf der Grundlage der Größe. Vorzugsweise ist der Größentrennschritt nicht-denaturierend, und die Nukleinsäurefragmente werden vor der Trennmethode denaturiert. Der Größentrennschritt kann z. B. mit Gelelektrophorese (z. B. Polyacrylamidgelelektrophorese) oder vorzugsweise HPLC durchgeführt werden. Die Markierungen werden anschließend von den aufgetrennten Fragmenten abgespalten, und anschließend werden die Markierungen mit der entsprechenden Nachweistechnologie nachgewiesen (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV-Spektrometrie).

#### 6. Restriktionskarten und RFLPs

[0268] Restriktionsendonukleasen erkennen kurze DNA-Sequenzen und schneiden DNA-Moleküle an diesen spezifischen Stellen. Einige Restriktionsenzyme (Rare-Cutters) schneiden DNA sehr selten, wodurch eine kleine Anzahl sehr großer Fragmente (mehrere tausend bis eine Million bp) erzeugt werden. Die meisten Enzyme schneiden DNA häufiger, wodurch eine große Anzahl kleiner Fragmente (weniger als einhundert bis mehr als eintausend bp) erzeugt wird. Im Mittel werden Restriktionsenzyme mit 4-Basen-Erkennungsstellen Stücke liefern, die 256 Basen lang sind, 6-Basen-Erkennungsstellen werden Stücke liefern, die 4.000 Basen lang sind, und 8-Basen-Erkennungsstellen werden Stücke liefern, die 64.000 Basen lang sind. Da hunderte von verschiedenen Restriktionsenzymen charakterisiert worden sind, kann DNA in viele verschiedene kleine Fragmente ge-

schnitten werden.

[0269] Eine breite Vielfalt von Techniken ist für die Analyse von DNA-Polymorphismen entwickelt worden. Das am weitesten verwendete Verfahren, der Restriktionsfragmentlängenpolymorphismus(RFLP)-Ansatz, kombiniert Restriktionsenzymverdau, Gelelektrophorese, Blotting auf eine Membran und Hybridisierung an eine klonierte DNA-Sonde. Polymorphismen werden als Variationen in den Längen der markierten Fragmente auf den Blots nachgewiesen. Der RFLP-Ansatz kann verwendet werden, um Basensubstitutionen zu analysieren, wenn die Sequenzveränderung in eine Restriktionsenzymstelle fällt, oder um Minisatelliten/(VNTRs) durch Auswahl von Restriktionsenzymen zu analysieren, die außerhalb der Wiederholungseinheiten schneiden. Die Agarosegele liefern üblicherweise nicht die notwendige Auflösung, um Minisatellit/VNTR-Allele zu unterscheiden, die sich um eine einzige Wiederholungseinheit unterscheiden, aber viele der Minisatelliten/VNTRs sind so variabel, daß dennoch hochinformativ Marker erhalten werden können.

[0270] In einer Ausführungsform der Erfindung werden Verfahren zum Bestimmen der Identität eines Nukleinsäuremoleküls oder zum Nachweisen eines selektierenden Nukleinsäuremoleküls in z. B. einer biologischen Probe unter Verwendung der Technik der Restriktionskartierung oder RFLPs bereitgestellt. Kurz gesagt umfassen solche Verfahren allgemein die Schritte: Erzeugen einer Reihe markierter Nukleinsäurefragmente, wobei die erzeugten Fragmente mit Restriktionsenzymen verdaut werden. Die markierten Fragmente werden erzeugt, indem ein Hybridisierungsschritt der markierten Sonden mit der verdauten Target-Nukleinsäure durchgeführt wird. Der Hybridisierungsschritt kann vor oder nach dem Restriktionsnukleaseverdau stattfinden. Die resultierenden verdauten Nukleinsäurefragmente werden dann nach Größe getrennt. Der Größentrennschritt kann z. B. mit Gelelektrophorese (z. B. Polyacrylamidgelelektrophorese) oder vorzugsweise HPLC durchgeführt werden. Die Markierungen werden anschließend von den getrennten Fragmenten abgespalten, und anschließend werden die Markierungen mit der entsprechenden Nachweisttechnologie nachgewiesen (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV-Spektrometrie).

## 7. DNA-Fingerprinting

[0271] DNA-Fingerprinting umfaßt das Display eines Satzes von DNA-Fragmenten von einer spezifischen DNA-Probe. Eine Vielzahl von DNA-Fingerprintingtechniken sind gegenwärtig verfügbar (Jeffreys et al., Nature 314: 67, 1985; Zabeua und Vos, 1992); "Selective Restriction Fragment Amplification: A General Method for DNA Fingerprinting," European Patent Application 92402629.7; Vos et al., "DNA FINGERPRINTING: A New Technique for DNA Fingerprinting." Nucl. Acids Res. 23: 4407-4414, 1996; Bates, S. R. E., Knorr, D. A., Weller, J. W., und Ziegele, J. S., "Instrumentation for Automated Molecular Marker Acquisition and Analysis." Kapitel 14, Seiten 239-255, in The Impact of Plant Molecular Genetics, edited by B. W. S. Sobral, veröffentlicht von Birkhauser, 1996.

[0272] Daher werden in einer Ausführungsform der Erfindung Verfahren zum Bestimmen der Identität eines Nukleinsäuremoleküls oder zum Nachweis eines selektierenden Nukleinsäuremoleküls in z. B. einer biologischen Probe unter Verwendung der Technik des DNA-Fingerprintings bereitgestellt. Kurz gesagt umfassen solche Verfahren allgemein die Schritte: Erzeugen einer Reihe markierter Nukleinsäurefragmente, gefolgt von Trennung der Fragmente nach Größe. Der Größentrennschritt kann z. B. mit Gelelektrophorese (z. B. Polyacrylamid-gelelektrophorese) oder vorzugsweise HPLC durchgeführt werden. Die Markierungen werden dann von den aufgetrennten Fragmenten abgespalten, und anschließend werden die Markierungen mit der entsprechenden Nachweisttechnologie nachgewiesen (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV-Spektrometrie).

[0273] Kurz gesagt, basiert DNA-Fingerprinting auf der selektiven PCR-Amplifikation von Restriktionsfragmenten aus einem Gesamtverdau von genomischer DNA: Die Technik schließt drei Schritte ein. 1) Restriktion der DNA-Fragmente und anschließende Ligation von Oligonukleotid-Adaptoren, 2) Die selektive Amplifikation von Sets von Restriktionsfragmenten, 3) Gelanalyse der amplifizierten Fragmente. Die PCR-Amplifikation der Restriktionsfragmente wird durch Verwendung der Adapter- und Restriktionsstellen-Sequenz als Zielstellen für die Primer Anhybridisierung erreicht. Die selektive Amplifikation wird durch die Verwendung von Primern erreicht, die sich in die Restriktionsfragmente hinein erstrecken und nur die Fragmente amplifizieren, in denen die Primerverlängerungen den Nukleotiden entsprechen, die die Restriktionsstellen flankieren.

[0274] Dieses Verfahren ergibt daher Sets von Restriktionsfragmenten, die durch eine Vielzahl von Verfahren visualisiert werden können (d. h. PAGE, HPLC oder andere Typen von Spektrometrie), ohne vorheriges Wissen der Nukleotidsequenz. Das Verfahren ermöglicht auch die Co-Amplifikation von großen Zahlen von Restriktionsfragmenten. Die Zahl von Fragmenten ist jedoch abhängig von der Auflösung des Nachweissystems. Typischerweise werden 50-100 Restriktionsfragmente amplifiziert und auf denaturierenden Polyacrylamidgele nachgewiesen. In der hier beschriebenen Anmeldung wird die Trennung durch HPLC durchgeführt.

[0275] Die DNA-Fingerprinting-Technik basiert auf der Amplifikation von Untergruppen von genomischen Restriktionsfragmenten unter der Verwendung von PCR. Die DNA wird mit Restriktionsenzymen geschnitten und Doppelstrang-Adapter werden an die Enden der DNA-Fragmente ligiert, um Templat-DNA für die einplifikati-

onsreaktionen zu erzeugen. Die Sequenz der ligierten Adapter und die angrenzenden Restriktionsenzyme (-stellen) dienen als Bindungsstellen für das DNA-Fingerprinting von Primern für die PCR-basierte Amplifikation. Selektive Nukleotide werden am 3'-Ende der PCR-Primer eingeschlossen, die dadurch nur eine DNA-Synthese von einer Untergruppe der Restriktionsstellen primen können. Nur Restriktionsfragmente, in denen die Restriktionsstelle flankierende Nukleotide zu den selektiven Nukleotiden passen, werden amplifiziert werden. Der DNA-Fingerprintprozess ergibt "Fingerprint"-Muster von verschiedenen Fragmentlängen, die charakteristisch und reproduzierbar für einen individuellen Organismus sind. Diese Fingerprints können dazu verwendet werden, sogar sehr eng verwandte Organismen zu unterscheiden, einschließlich nahezu isogener Linien. Die Unterschiede in den Fragmentlängen können auf Basenveränderungen in der Restriktionsstelle oder der Primerverlängerungsstelle oder auf Insertionen oder Deletionen im Körper des DNA-Fragments zurückgeführt werden.

[0276] Die Abhängigkeit von dem Wissen um die Sequenz des Zielgenoms wird durch die Verwendung von Adaptoren von bekannter Sequenz eliminiert, die an die Restriktionsfragmente ligiert werden. Die PCR-Primer sind für die bekannten Sequenzen der Adaptoren und der Restriktionsstellen spezifisch. Die Schritte des genetischen Fingerprintingverfahrens sind unten beschrieben.

1) Restriktion und Ligation. Die Restriktionsfragmente von genomischer DNA werden durch Verwendung zweier verschiedener Restriktionsenzymen erzeugt: ein selten schneidendes Enzym (das sechs-Base-erkennende Enzym EcoRI) und ein häufig schneidendes Enzym (das vier-Basen-erkennende Enzym MseI). Drei verschiedene Typen von Fragmenten werden produziert: einige mit EcoRI-Schnitten an beiden Enden, einige MseI-Schnitten an beiden Enden und einige mit einem EcoRI-Schnitt an einem Ende und mit einem MseI-Schnitt am anderen Ende. Doppelsträngige Adaptoren werden dann an die überstehenden Enden der DNA-Fragmente ligiert, was Templat-DNA für die Amplifikation erzeugt. Die Adaptoren sind spezifisch für entweder die EcoRI-Stelle oder die MseI-Stelle. Die Restriktion und Ligation findet in einer einzelnen Reaktion statt. Die Ligation des Adaptors an die verdaut DNA verändert die Restriktionsstelle, um so zu verhindern, daß eine zweite Restriktion stattfindet, nachdem die Ligation stattgefunden hat.

2) Präselektive Amplifikation. Die Sequenzen der Adaptoren und Restriktionsstellen dienen als Primerbindungsstellen für die "präselektive PCR-Amplifikation". Die präselektiven Primer weisen jeder ein "selektives" Nukleotid auf, das das Untersetz von Restriktionsfragmenten erkennen wird, die das passenden Nukleotid stromabwärts von der Restriktionsstelle aufweisen. Die primären Produkte der präselektiven PCR sind diejenigen Fragmente, die einen MseI-Schnitt und EcoRI-Schnitt aufweisen und auch das passende interne Nukleotid aufweisen. Die präselektive Amplifikation erreicht eine 16-fache Verringerung der Komplexität des Fragmentgemisches.

3) Selektive Amplifikation mit CMST-markierten Primern. Die Komplexität des PCR-Produkts wird weiter reduziert (256-fach) und Fragmente werden mit einem Set von CMSTs durch Durchführen einer zweiten PCR unter der Verwendung von selektiven Primern, markiert mit CMSTs, markiert. Es ist möglich, aus unter 64 verschiedenen Primerpaaren (die sich aus allen möglichen Kombinationen von acht MseI- und acht EcoRI-Primern ergeben) für die Amplifikation auszuwählen. Jeder dieser Primer besitzt drei selektive Nukleotide. Das erste ist dasselbe, wie das in der präselektiven Amplifikation verwendete, die anderen können jede von den 16 möglichen Kombinationen der vier Nukleotide sein. Nur die Untergruppe von Fragmenten, die passende Nukleotide an allen drei Positionen aufweist, wird zu dieser Phase in der Amplifikation amplifiziert werden.

## 8. Anwendung spaltbarer Marker auf die Genotypisierung und den Polymorphismus-Nachweis

### a. Einführung

[0277] Obgleich einige bekannte Human-DNA-Polymorphismen auf Insertionen, Deletionen oder anderen Umordnungen nicht-wiederholter Sequenzen beruhen, beruht die große Mehrzahl auf entweder Einzelbasensubstitutionen oder auf Variationen in der Anzahl von Tandemwiederholungen. Basensubstitutionen sind im menschlichen Genom reichlich vorhanden, wobei sie im Mittel einmal alle 200–500 bp auftreten. Längenvariationen in Blöcken von Tandemwiederholungen sind ebenfalls im Genom üblich, mit wenigstens einigen 10.000 eingestreuten polymorphen Stellen (Loci genannt). Wiederholungslängen für Tandem-Wiederholungs-Polymorphismen reichen von 1 bp in  $(dA)_n(dT)_n$ -Sequenzen bis wenigstens 170 bp in  $\alpha$ -Satelliten-DNA. Tandem-Wiederholungs-Polymorphismen können in zwei große Gruppen unterteilt werden, die aus Minisatelliten/variable Anzahl von Tandem-Wiederholungen (VNTRs), mit typischen Wiederholungslängen von einigen zehn Basenpaaren und mit einigen 10.000 Gesamtwiederholungseinheiten, und Mikrosatelliten, mit Wiederholungslängen von bis zu 6 bp und mit maximalen Gesamtlängen von etwa 70 bp, bestehen. Die meisten Mikrosatelliten-Polymorphismen, die bisher identifiziert worden sind, haben auf  $(dC-dA)_n$ - oder  $(dG-dT)_n$ -Dinukleotidwiederholungssequenzen beruht. Die Analyse von Mikrosatelliten-Polymorphismen umfaßt die Amplifikation mit der Polymerasekettenreaktion (PCR) eines kleinen DNA-Fragments, das einen Block von Wiederholungen enthält, gefolgt von Elektrophorese der amplifizierten DNA auf denaturierendem Polyacrylamidgel. Die

PCR-Primer sind komplementär zu einzigartigen Sequenzen, die die Wiederholungsblöcke flankieren. Polyacrylamidgele, anstelle von Agarosegelen, werden traditionellerweise für Mikrosatelliten verwendet, weil die Allele sich in der Größe oft nur um eine einzige Wiederholung unterscheiden.

[0278] So werden in einem Aspekt der vorliegenden Erfindung Verfahren zum Genotypisieren eines ausgewählten Organismus zur Verfügung gestellt, die die Schritte umfassen: (a) Erzeugen markierter Nukleinsäuremoleküle aus einem ausgewählten Targetmolekül, wobei eine Markierung mit einem bestimmten Fragment korrelativ ist und mit Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie nachgewiesen werden kann, (b) Abtrennen der markierten Moleküle nach Sequenzlänge, (c) Abspalten der Markierung vom markierten Molekül und (d) Nachweisen der Markierung mit Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie und Bestimmen des Genotyps des Organismus daraus.

[0279] In einem weiteren Aspekt werden Verfahren zum Genotypisieren eines ausgewählten Organismus zur Verfügung gestellt, die die Schritte umfassen: (a) Kombinieren eines markierten Nukleinsäuremoleküls mit einem ausgewählten Targetmolekül unter Bedingungen und für einen Zeitraum, die ausreichend sind, um Hybridisierung des markierten Moleküls an das Targetmolekül zu ermöglichen, wobei eine Markierung mit einem bestimmten Fragment korrelativ ist und mit Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie nachgewiesen werden kann, (b) Abtrennen der markierten Fragmente nach Sequenzlänge, (c) Abspalten der Markierung vom markierten Fragment und (d) Nachweisen der Markierung mit Nicht-Fluoreszenz-Spektrometrie oder Potentiometrie und Bestimmen des Genotyps des Organismus daraus.

#### b. Anwendung spaltbarer Markierungen auf die Genotypisierung

[0280] Ein PCR-Ansatz, um Restriktionsfragmentlängenpolymorphismus (RFPL) zu identifizieren, kombiniert Gelelektrophorese und den Nachweis von Markierungen, die mit spezifischen PCR-Primern assoziiert sind. Im allgemeinen wird ein PCR-Primer eine spezifische Markierung aufweisen. Die Markierung wird daher einen Satz von PCR-Primern repräsentieren und daher eine vorbestimmte DNA-Fragmentlänge. Polymorphismen werden als Variationen in der Länge der markierten Fragmente in einem Gel oder aus einem Gel eluierend nachgewiesen. Polyacrylamidgelelektrophorese wird üblicherweise die notwendige Auflösung liefern, um Minisatellit/VNTR-Allele zu unterscheiden, die sich um eine einzige Wiederholungseinheit unterscheiden. Die Analyse von Mikrosatellitenpolymorphismen umfaßt die Amplifikation mit der Polymerasekettenreaktion (PCR) eines kleinen DNA-Fragments, das einen Wiederholungsblock enthält, gefolgt von der Elektrophorese der amplifizierten DNA auf denaturierendem Polyacrylamidgel oder gefolgt von der Trennung von DNA-Fragmenten mit HPLC. Die amplifizierte DNA wird unter Verwendung von Primern markiert werden, die spaltbare Markierungen am 5'-Ende des Primers aufweisen. Die Primer werden in die neu synthetisierten Stränge durch Kettenverlängerung eingebaut. Die PCR-Primer sind komplementär zu einzigartigen Sequenzen, die die Wiederholungsblöcke flankieren. Minisatelliten/VNTR-Polymorphismen können ebenfalls amplifiziert werden, weitgehend so wie bei den oben beschriebenen Mikrosatelliten.

[0281] Beschreibungen vieler Arten von DNA-Sequenzpolymorphismen haben die grundlegende Basis für das Verständnis der Struktur des menschlichen Genoms geliefert (Botstein et al., *Am. J. Human Genetics* 32: S. 314, 1980; Donis-Keller, *Cell* 51: 319, 1987; Weissenbach et al., *Nature* 359: 794). Die Konstruktion umfangreicher Kopplungskarten-Systeme ist durch die Verwendung dieser DNA-Polymorphismen erleichtert worden und hat ein praktisches Mittel zur Lokalisierung von Krankheitsgenen durch Kopplung geliefert. Mikrosatelliten-Dinukleotidmarker haben sich als sehr leistungsfähige Tools bei der Identifizierung menschlicher Gene erwiesen, bei denen sich gezeigt hat, daß sie Mutationen enthalten und in einigen Fällen Krankheiten verursachen. Genomische Dinukleotidwiederholungen sind hochpolymorph (Weber, 1990, *Genomic Analysis*, Vol. 1, S. 159–181, Cold Spring Laboratory Press, Cold Spring Harbor, NY; Weber und Wong, 1993, *Hum. Mol. Genetics*, 2, S. 1123) und können bis zu 24 Allele besitzen. Mikrosatelliten-Dinukleotidwiederholungen können mit PCR unter Verwendung von Primern amplifiziert werden, die zu den einzigartigen Regionen komplementär sind, die das Dinukleotid umgeben. Im Anschluß an die Amplifikation können mehrere amplifizierte Loci vor einem Größentrennschritt kombiniert (vervielfältigt) werden. Das Verfahren der Anwendung der amplifizierten Mikrosatellitenfragmente auf einen Größentrennschritt und des anschließenden Identifizierens der Größe und daher des Allels ist als Genotypisierung bekannt. Über chromosomenspezifische Marker, die ein hohes Vervielfältigungsniveau erlauben, ist zum Durchführen vollständiger Genomscans für Kopplungsanalyse berichtet worden (Davies et al., 1994, *Nature*, 371, S. 130).

[0282] Markierungen können mit großer Wirkung beim Genotypisieren mit Mikrosatelliten verwendet werden. Kurz gesagt, werden die PCR-Primer konstruiert, um Markierungen zu tragen, und in einer sorgfältig ausgewählten PCR-Reaktion verwendet, um Di-, Tri- oder Tetranukleotid-Wiederholungen zu amplifizieren. Die Amplifikationsprodukte werden anschließend nach Größe mit Verfahren, wie etwa HPLC oder PAGE getrennt. Die DNA-Fragmente werden anschließend in einer temporären Art und Weise gesammelt, die Markierungen von ihren entsprechenden DNA-Fragmenten abgespalten und die Länge aus einem Vergleich mit internen Standards im Größentrennschritt abgeleitet. Allel-Identifizierung wird aus der Bezugnahme zur Größe der amplifi-

zierten Produkte durchgeführt.

[0283] Beim Ansatz mit spaltbaren Markierungen für die Genotypisierung ist es möglich, mehrere Proben in einem einzigen Trennschritt zu kombinieren. Es gibt zwei allgemeine Wege, auf denen dies durchgeführt werden kann. Das erste allgemeine Verfahren für ein Screening mit hohem Durchsatz ist der Nachweis eines einzelnen Polymorphismus in einer großen Gruppe von Individuen. In diesem Szenario wird ein einzelner oder „nested“ Satz von PCR-Primern verwendet, und jede Amplifikation wird mit einem DNA-Probentyp pro Reaktion durchgeführt. Die Anzahl von Proben, die im Trennschritt kombiniert werden können, ist proportional zur Anzahl spaltbarer Markierungen, die pro Nachweisthechnologie erzeugt werden können (d. h. 400–600 für Massenspektrometrie-Markierungen). Es ist daher möglich, 1 bis mehrere Polymorphismen einer großen Gruppe von Individuen gleichzeitig zu identifizieren. Der zweite Ansatz ist, mehrere Sätze von PCR-Primern zu verwenden, die zahlreiche Polymorphismen auf einer einzigen DNA-Probe identifizieren können (ein Individuum z. B. genotypisieren können). Bei diesem Ansatz werden PCR-Primer in einer einzigen Amplifikationsreaktion kombiniert, die PCR-Produkte unterschiedlicher Länge erzeugt. Jedes Primerpaar oder jeder eingebaute Satz ist durch eine spezifische spaltbare Markierung codiert, was impliziert, daß jedes PCR-Fragment mit einer spezifischen Markierung codiert sein wird. Die Reaktion wird in einem einzigen Trennschritt durchgeführt (siehe unten). Die Anzahl von Proben, die im Trennschritt kombiniert werden können, ist proportional zur Anzahl spaltbarer Markierungen, die pro Nachweisthechnologie erzeugt werden können (d. h. 400–600 für Massenspektrometer-Markierungen).

[0284] Die Genotypisierung kann auch auf landwirtschaftliche Proben angewendet werden. Zum Beispiel ermöglicht die Amplifizierter Polymorphismus Längen Polymorphismus (AFLP)-Analyse eine landwirtschaftlich bedeutsame Gruppierung des Keimplasmas. Die Zusammenfassung von QTL-Effekten kann verschiedene Regionen des Genoms ergeben, die konsistent wichtige Determinanten von vielen quantitativen Eigenschaften sind. Die Assoziation von AFLP- und QTL-Polymorphismen sollte von prädiktiver Brauchbarkeit beim Erstellen von Paarungsgruppen sein. Weiterhin sollte das Wissen im Hinblick auf die genomische Architektur von verwandten und nicht verwandten Keimplasmen in Schlüsselregionen des Genoms eine systematische Trennung auf Basis von Selektionsantworten in verwandten Keimplasmen ermöglichen. Es sollte auch die systematische Introgression von exotischen Keimplasmen erleichtern. Zuletzt sollte eine übergreifende Beschreibung von Keimplasma auf dem molekularen Niveau, gekoppelt mit extensiver QTL-Information, es Züchtern ermöglichen, eine mittlere Performance beizubehalten und gleichzeitig genetische Diversität zu erreichen.

[0285] Sobald Werkzeuge für die Routinecharakterisierung zur Verwendung von genetischen Ressourcen erhältlich sind, wird es einen größeren Impetus für die Konservierung von genetischen Ressourcen geben.

[0286] Lander und Botstein (Lander, E. S. und D. Botstein, "Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps" *Genetics* 121: 185–189, 1987, beschrieben die Entwicklung von Gedanken, die zum Abgleich der Mendelschen Theorie der bestimmten Vererbung und der Beobachtung führten, daß viele Marker eine kontinuierliche Variation zeigen. Die Entwicklung von molekularen Markertechnologien wird die Entwicklung von ausreichend dichten Karten ermöglichen, um die Abtrennung von quantitativen Markern zu initiieren. Ein hauptsächlicher Impetus für den QTL-Nachweis ist es, die zugrunde liegenden Determinanten in einem angewandten Züchterkontext zu manipulieren. Parterson et al. "DNA markers in plant improvement" *Adv. in Agron.* 46: 39–90, 1991; und Dedley, J. W. "Comparison of genetic distance estimators using molecular marker data", *Proc. Second Plant Breeding Symposium of the Crop Sci. Soc. Amer. and Amer. Soc. Hort.*, Corvallis, OR, Amer. Soc. Hort. Sci. Alexandria, VA, 1994, stellten alle ausgezeichnete Überblicke über die potentiellen Anwendungen dieser Techniken zur Verfügung. Zum Beispiel kann die Assoziation von quantitativer Markerexpression mit kartierten Markern als eine Basis für Entscheidungen im Hinblick auf die Manipulation von Züchtungspopulationen, für die molekulare Marker unterstützte Selektion (MMAS) oder zur Entdeckung von wünschenswerten Genen in exotischen Keimplasmen (siehe auch Tanksley et al. "Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines", *Theor. Appl. Genet.* 92: 191–203, 1996) führen. QTL für einen Bereich von Markern, z. B. Ausbeute, Malzqualität, Winterhärte, und Erkrankungsresistenz, wurden in einer Zahl von Gerste Keimplasmaquellen lokalisiert (zusammengefaßt von Hayes et al., "Barley genome mapping and its applications" in P. P. Jauhar (ed.) *Methods for Genome Analysis in Plants*, CRC Press, Boca Raton, USA, 1996; und Hayes, P. M. et al., "Multiple disease resistance loci and their relationship to agronomic and quality loci in a spring barley population", *JQTL*, 1996, see <http://probe.nalusda.gov:8000/otherdocs/jqtl/index.htm>). Diese Berichte, wie diejenigen in anderen Getreidsorten, waren größtenteils beschreibend.

[0287] Die genetische Basis dieser QTL, obwohl sie von größerem theoretischem Interesse und signifikanter praktischer Wichtigkeit sind, ist im allgemeinen nicht bekannt. Es wurde vorgeschlagen, daß QTL Konsequenzen von Allelen an Loci sind, wo die Effekte von anderen Allelen auf eine einfache Mendelsche Weise interpretiert werden können. Falls das der Fall wäre, dann sollten sich QTL-Allele in einer additiven Weise verhalten und sie sollten von einem Set von Keimplasma zu einem anderen transferierbar sein. Alternativ können komplexe Typen die Endergebnisse von komplexen Signalwegen sein (siehe z. B. Dawkins, R. "The extended phenotype" Oxford University Press, Oxford, UK, 1982). Falls dem so ist, stellt QTL die Effekte von Allelen dar, die

den komplexen Signalweg stören, der in quantitativen Phänotypen gipfelt. In diesem Szenario wäre es einfach, den quantitativen Phänotyp zu verändern (d. h. die Anwesenheit eines nicht wünschenswerten Allels an jedem Punkt in dem Signalweg wäre ausreichend, den Signalweg zum Entgleisen zu bringen), jedoch schwierig den quantitativen Phänotyp zu erhalten (d. h. ein Wirt mit gewünschten Allelen an multiplen Loci würde erforderlich sein, um den quantitativen Phänotyp zu erhalten). Falls dies der Fall wäre, dann wäre die Introgression von QTL-Allelen in nicht verwandtes Keimplasma ein relativer Erfolg- oder Fehlschlagansatz: Keimplasma, das nicht mit der kartierenden Referenz-Population verwandt wäre, könnte gut nicht wünschenswerte Allele an Loci aufweisen, wo vorteilhafte Allele in der kartierenden Referenz-Population fixiert wären.

[0288] Die Komplexität der Kontrolle wird möglicherweise mit der untersuchten Eigenschaft variieren. Zuletzt sollte sogar der komplexeste Phänotyp letztendlich auf eine Serie von diskreten, wenn auch hoch interaktiven, Schritten reduzierbar sein. Relativ wenige QTL sind für komplexe Phänotypen nachgewiesen, wie zum Beispiel Ausbeute, Komponenten der Malzqualität und die quantitative Resistenz gegenüber biotischen und abiotischen Stressfaktoren (zusammengefaßt durch Hayes et al., "Barley genome mapping and its applications" in P. P. Jauhar (ed.) *Methods of Genome Analysis in Plants*: CRC Press, Boca Raton, USA, 1996). In einigen Fällen können Kandidatengene als QTL-Determinanten angegeben werden: Schüttelresistenz als eine Determinante der Ausbeute (Hayes et al., 1993, wie oben) und hydrolytische Enzyme als Determinanten von Malzextrakt (die Menge von löslichen Kohlenhydraten, die als ein Substrat für die Fermentation erhältlich sind) (siehe Hayes, P. M., 1996m, oben). Barley genome mapping: new insights into the malting quality of the world's oldest crop. (siehe, z. B. MBAA 33: 223-225).

[0289] In anderen Fällen können bis jetzt keine Kandidatengene als Determinanten von komplexen Phänotypen angegeben werden. Ein Beispiel ist die quantitative adulte Pflanzenresistenz gegenüber Streifenrost, wobei Resistenz-QTL nicht mit den berichteten Kartenlokationen von Resistenzgenen zusammenfallen, die Muster von Mendelscher Vererbung zeigen (siehe Hayes, P. M. et al., "Multiple disease resistance loci and their relationship to agronomic and quality loci in a spring barley population", JQTL, 1996 und <http://probe.nalus-da.gov:8000/otherdocs/jqtl/index.htm>).

[0290] DNA-Niveau-Polymorphismen können auch dazu verwendet werden, Themen der genetischen Diversität zu untersuchen. Genetische Diversität wurde in einer Zahl von Gerste-Keimplasma-Arrays gemessen, unter der Verwendung von Werkzeugen, die von Abstammungsanalyse (siehe, z. B. Eslick, R. F. et al., "Genetic engineering as a key to ater use efficiency", *Agric. Meteor.* 14: 13-23, 1974), über morphologische Marker (siehe, z. B., Tolbert, D. M. et al., "A diversity analysis of a world collection of barley" *Crop. Sci.* 19: 789-794, 1979) bis zu molekularen Markern verschiedener Sorten (Melchinger, A. E. et al. "Relationships among European barley germplasm: I. Genetic diversity among winter and spring cultivars revealed by RFLPs" *Crop Sci.* 34: 1191-1199, 1994; Saghai-Maroo, M. A. et al. "RFLPs in cultivated barley and their application in the evaluation of malting quality cultivars" *Hereditas* 121: 2129, 1994) reichten. Molekulare Marker sind aus dieser Hinsicht besonders attraktiv, da sie reichlich vorhandene adaptiv neutrale Referenzpunkte zur Verfügung stellen. Im Fall von molekularen Polymorphismen waren die Zeit, die Ressourcen die erforderlich waren, die Daten zu erzeugen, eine Begrenzung.

[0291] Die Kopplungs-Karten-Konstruktion ist der erste Schritt in einer systematischen QTL-Analyse. Werkzeuge, wie zum Beispiel die CMST-Technologie-Plattform wie hier bei der Kartierung verwendet beschrieben, sollten auch dazu brauchbar sein es den Züchtern zu ermöglichen, die genetische Diversität auszudehnen ohne den Wert des Arbeits-Keimplasmapools zu verringern. Bis jetzt war nicht ausreichend Information erhältlich, um eine Charakterisierung der Diversität mit den Ergebnissen von QTL-Analysen zu integrieren. Der Schlüssel für solch eine integrative Strategie ist eine Markertechnologie, die schnell und Kosten-effektiv ist und ausreichend Polymorphismen über das Genom hinweg zur Verfügung stellt. Die vorliegende Erfindung stellt diesen Schlüssel zur Verfügung.

#### c. Enzymatischer Mutationsnachweis und die Anwendungen von Markierungen.

[0292] In dieser besonderen Anwendung oder diesem besonderen Verfahren werden Fehlpaarungen in Heteroduplexen durch enzymatische Spaltung von Fehlpaarungs-Basenpaaren in einer gegebenen Nukleinsäureduplex nachgewiesen. DNA-Sequenzen, die auf das Vorhandensein von einer Mutation getestet werden sollen, werden mit PCR unter Verwendung eines spezifischen Satzes von Primern amplifiziert, die amplifizierten Produkte werden denaturiert und mit denaturierten Bezugsfragmenten vermischt und hybridisiert, was zur Bildung von Heteroduplexen führt. Die Heteroduplexen werden anschließend mit Enzymen behandelt, die die Duplex erkennen und spalten, wenn eine Fehlpaarung vorhanden ist. Solche Enzyme sind Nuklease S1, Mungobohnen-Nuklease, "Resolvasen", T4-Endonuklease IV, etc.. Im wesentlichen kann jedes Enzym verwendet werden, das Fehlpaarungen in vitro erkennt und die resultierende Fehlpaarung spaltet. Nach Behandlung mit dem geeigneten Enzym werden die Duplexe nach Größe mit z. B. HPLC oder PAGE getrennt. Die DNA-Fragmente werden temporär gesammelt. Markierungen werden abgespalten und nachgewiesen. Das Vorhandensein einer Mutation wird durch die Verschiebung der Mobilität des Fragments relativ zu einem Wildtyp-Be-

zugsfragment nachgewiesen.

#### d. Anwendungen von Markierungen auf dem Oligonukleotid-Ligationsassay (OLA).

[0293] Der Oligonukleotid-Ligationsassay, wie er ursprünglich von Landegren et al. (Landegren et al., Science 241: 487, 1988) beschrieben ist, ist eine nützliche Technik zur Identifizierung von (bekannten) Sequenzen in sehr großen und komplexen Genomen. Das Prinzip der OLA-Reaktion beruht auf der Fähigkeit von Ligase, zwei diagnostische Oligonukleotide kovalent zu verknüpfen, wenn sie benachbart zueinander auf einem gegebenen DNA-Target hybridisieren. Wenn die Sequenzen an der Sondenverbindungsstelle nicht perfekt basen-gepaart sind, werden die Sonden nicht von der Ligase verknüpft werden. Die Fähigkeit einer thermostabilen Ligase, potentielle Einzelbasenpaarunterschiede zu unterscheiden, wenn sie am 3'-Ende der "stromaufwärtigen" Sonde angeordnet sind, liefert die Möglichkeit zur Einzelbasenpaarauflösung (Barony, PNAS USA 88: 189, 1991). Bei der Anwendung von Markierungen würden die Markierungen an die Sonde gebunden werden, die an das amplifizierte Produkt ligiert wird. Nach Abschluß der OLA werden die Fragmente auf der Grundlage ihrer Größe getrennt, die Markierungen abgespalten und mit Massenspektrometrie nachgewiesen.

#### e. Sequenzspezifische Amplifikation

[0294] PCR-Primer mit einem 3'-Ende, das entweder zu einer Mutanten- oder normalen Oligonukleotidsequenz komplementär ist, können verwendet werden, um das eine oder das andere Allel selektiv zu amplifizieren (Newton et al., Nuc. Acids Res., 17, S. 2503; et al., 1989, Genomics, 5, S. 535; Okayama et al., 1989, J. Lab. Clin. Med., 114, S. 105; Sommer et al., 1989, Mayo Clin. Proc., 64, 1361; Wu et al., PNAS USA, 86, S. 2757). Üblicherweise werden die PCR-Produkte nach der Amplifikation mit PAGE sichtbar gemacht, aber das Prinzip sequenzspezifischer Amplifikation kann auf Festphasenformate angewendet werden.

#### f. Potentielle Anwendung von Markierungen auf einige Assays auf Amplifikationsbasis.

[0295] Genotypisierung von Viren: Eine potentielle Anwendung von Markierungen ist die Genotypisierung oder Identifizierung von Viren durch Hybridisierung mit markierten Sonden. F+-RNA-Coliphagen können z. B. nützliche Kandidaten als Indikatoren für Enteroviruskontamination sein. Genotypisierung mit Nukleinsäurehybridisierungsverfahren ist eine verlässliche, schnelle, einfache und preiswerte Alternative zur Serotypisierung (Kafatos et al., Nucleic Acids Res. 7: 1541, 1979). Amplifikationstechniken und Nukleinsäurehybridisierungstechniken sind mit Erfolg verwendet worden, um eine Vielzahl von Mikroorganismen zu klassifizieren, einschließlich E. coli (Feng, Mol. Cell Probes 7: 151, 1993), Rotavirus (Sethabutr et al., J. Med. Virol. 37: 192, 1992), Hepatitis-C-Virus (Stuyver et al., J. Gen Virol. 74: 1093, 1993) und Herpes-simplex-Virus (Matsumoto et al., J. Virol. Methods 40: 119, 1992).

[0296] Prognostische Anwendungen von Mutationsanalyse bei Krebserkrankungen: Genetische Veränderungen sind in einer Vielzahl experimenteller Säuger- und Human-Neoplasmen beschrieben worden und stellen die morphologische Grundlage für die Abfolge morphologischer Veränderungen dar, die bei Karzinogenese beobachtet wird (Vogelstein et al., NEJM 319: 525, 1988). In den letzten Jahren, mit dem Aufkommen von molekularbiologischen Techniken, sind Allelverluste auf bestimmten Chromosomen oder Mutation von Tumorsuppressorgenen sowie Mutationen in mehreren Onkogenen (z. B., c-myc, c-jun und die ras-Familie) die am häufigsten untersuchten Einheiten gewesen.

[0297] Frühere Arbeiten (Finkelstein et al., Arch Surg. 128: 526, 1993) haben eine Korrelation zwischen spezifischen Typen von Punktmutationen im K-ras-Onkogen und dem Stadium bei Diagnose bei kolorektalem Karzinom identifiziert. Die Ergebnisse legen nahe, daß die Mutationsanalyse wichtige Informationen über Tumoraggressivität liefern könnte, einschließlich dem Muster und der Verteilung von Metastasen. Der prognostische Werte von TP53- und K-ras-2-Mutationsanalyse bei Stufe III-Karzinomen des Colons ist vor kurzem nachgewiesen worden (Pricolo et al., Am. J. Surg. 171: 41, 1996). Es ist daher offensichtlich, daß die Genotypisierung von Tumoren und präkanzerösen Zellen und spezifischer Mutationsnachweis bei Behandlung von Krebserkrankungen bei Menschen zunehmend wichtig werden wird.

#### 9. Einzel-Nukleotid Verlängerungsassay

[0298] Die Primerverlängerungstechnik für den Nachweis von Einzelnukleotiden in genomischer DNA wurde erst durch Sokolov 1989 beschrieben (Nucleic Acids Res. 18(12): 3671, 1989). In diesem Artikel beschrieb Sokolov die Einzelnukleotidverlängerung von 30-Meren und 20-Meren, die zu der bekannten Sequenz des zystischen Fibrosegens komplementär waren. Es wurde gezeigt, daß das Verfahren die Fähigkeit aufwies, konkret eine einzelne Nukleotidveränderung innerhalb des Gens zu identifizieren. Das Verfahren basierte auf der Verwendung von radiomarkierten Desoxynukleotiden für ein Markierungsverfahren in dem Einzelnukleotidverlän-



gerungsassay.

[0299] Spätere Publikationen beschrieben die Verwendung von Einzelnukleotidverlängerungsassays für genetische Erkrankungen, wie zum Beispiel Hämophilie B (Faktor IX) und das zystische Fibrosegen (siehe, z. B., Kuppaswamy et al., PNAS USA 88: S. 1143–1147, 1991). Kuppaswamy et al. zeigten, daß der Einzelnukleotidverlängerungsassay dazu verwendet werden könnte, um genetische Erkrankungen nachzuweisen, wobei die Anwendung der Nachweis von Hämophilie B (Faktor IX) und das zystische Fibrosegen waren. Wiederum basiert dieses Verfahren auf dem einzelnen Nukleotidprimer, der an eine Sequenz hybridisiert wird, die an einen bekannten Single Nucleotide Polymorphismus angrenzt. Die geprimte genomische DNA wird dann Bedingungen unterzogen, in denen Taq-Polymerase ein  $^{32}\text{P}$ -markiertes dNTP hinzufügen wird, falls die Stelle hinter der Stelle von Interesse komplementär zu dem alpha markierten Reaktionsgemisch ist.

[0300] Kürzlich wurden die Parameter des Einzelnukleotidverlängerungsassays im Hinblick auf den quantitativen Bereich, die Variabilität und Multiplex-Analyse genau beschrieben. Daher beschrieben 1996 Greenwood und Burke (Greenwood, A. D. und Burke, D. T. (1996) Genome Research, 6, S. 336–348) genau die Parameter des Einzelnukleotidverlängerungsassays im Hinblick auf den quantitativen Bereich, die Variabilität und die Multiplex-Analyse. RNA diente als ein Templat für die PCR-Amplifikation von einer Sequenz von Interesse, die einen einzel-Base Unterschied zwischen zwei Allelen aufweist. Jedes PCR-erzeugte Templat wird auf die Anwesenheit, Abwesenheit oder relative Menge jedes einzelnen Allels durch Annealen eines Primers analysiert, der eine Base 5' zu dem Polymorphismus liegt und sich um eine markierte Base verlängert (oder eine markierte Base verwendet). Nur wenn die korrekte Base in der Reaktion erhältlich ist, wird ein Einbau am 3'-Ende des Primers auftreten. Die Verlängerungsprodukte werden dann analysiert (herkömmlich durch PAGE).

[0301] Daher basiert diese Strategie auf der Genauigkeit der DNA-Polymerase, nur das korrekt gepaarte Nukleotid an das 3'-Ende des Templat-hybridisierten Primers anzufügen. Da nur ein Didesoxy-Terminatornukleotid pro Reaktion hinzugefügt wird, ist dies eine einfache Weise herauszubekommen, welcher Primer in allen vier Typen von dNTPs verlängert wurde.

[0302] Daher werden innerhalb eines Aspekts der Erfindung Verfahren zum Nachweis einer einzelnen ausgewählten Nukleotids innerhalb eines Nukleinsäuremoleküls zur Verfügung gestellt, umfassend die Schritte von (a) Hybridisieren in mindestens zwei getrennten Reaktionen einen markierten Primer und ein Zielnukleinsäuremolekül unter Bedingungen und für eine Zeitdauer, die ausreichend ist, eine Hybridisierung des Primers an das Zielnukleinsäuremolekül zu ermöglichen, wobei jede Reaktion ein Enzym enthält, das einen Nukleotidkettenterminator hinzufügen wird, und einen Nukleinkettenterminator, komplementär zu Adenosin, Cytosin, Guanodin, Thymin oder Uracil, und wobei jede Reaktion einen verschiedenen Nukleotidkettenterminator enthält, (b) Auftrennen von markierten Primern mittels Größe, (c) Abspalten des Markers von dem markierten Primer und (d) Nachweisen des Markers durch nicht-fluoreszente Spektrometrie oder Potentiometrie und daraus Nachweisen der Anwesenheit des ausgewählten Nukleotids innerhalb des Nukleinsäuremoleküls.

[0303] Wie hier angegeben, kann eine große Vielzahl von Abtrennverfahren verwendet werden, einschließlich, z. B. Flüssigkeitschromatographie-Mitteln, wie zum Beispiel HPLC. Zusätzlich kann eine große Vielzahl von Nachweisverfahren verwendet werden, einschließlich, zum Beispiel, Massenspektrometrie, Infrarotspektrometrie, Ultraviolett-Spektrometrie oder potentiostatische Amperometrie. Auch können mehrere verschiedene Enzyme verwendet werden (z. B. eine Polymerase) sowie jede der hier zur Verfügung gestellten Markierungen. Innerhalb bestimmter bevorzugter Ausführungsformen weist jeder Primer, der innerhalb einer Reaktion verwendet wird, einen verschiedenen einmaligen Marker auf. Auf diese Weise können multiple Proben (oder multiple Stellen) gleichzeitig auf die Anwesenheit von ausgewählten Nukleotiden hin getestet werden.

[0304] Einzelnukleotidtests, so wie diejenigen hier beschrieben, können dazu verwendet werden, um polymorphe Varianten nachzuweisen oder eine biologische Probe auf die Anwesenheit eines spezifischen Nukleotids innerhalb oder nahe einer bekannten Sequenz hin zu untersuchen. Die Zielnukleinsäuremoleküle schließen nicht nur DNA ein (z. B. genomische DNA), sondern auch RNA.

[0305] Im allgemeinen schießt dieses Verfahren das Hybridisieren eines Primers an die Ziel-DNA-Sequenz auf solche Weise ein, daß das 3'-Ende des Primers unmittelbar an die Mutation, die nachgewiesen und identifiziert werden soll, angrenzt. Das Verfahren ist ähnlich zu der Sanger Sequenzreaktion, mit der Ausnahme, daß nur das Didesoxynukleotid eines bestimmten Nukleotids zum Reaktionsgemisch hinzugefügt wird. Jedes Didesoxynukleotid wird mit einem einmaligen Marker markiert. Von den vier Reaktionsgemischen wird nur eines einen Didesoxyterminator an die Primersequenz anfügen. Falls die Mutation vorhanden ist, wird sie durch den einmaligen Marker auf dem Didesoxynukleotid nachgewiesen werden und seine Identität etabliert werden. Auch können multiple Mutationen durch Markieren des DNA-Primers mit einem einmaligen Marker gleichzeitig ermittelt werden. Innerhalb eines Aspekts der Erfindung werden Verfahren zur Analyse von Einzelnukleotid-Mutationen aus einer bestimmten biologischen Probe zur Verfügung gestellt, umfassend die Schritte von Exponieren von Nukleinsäuren aus einer biologischen Probe und Kombinieren der ausgesetzten Nukleinsäuren mit einer oder mehreren ausgewählten Nukleinsäuren, die markiert oder nicht markiert sein können, unter Bedingungen und für eine Zeitdauer, die ausreichend ist, für die Sonden an die Nukleinsäuren zu hybridisieren, wobei der Marker, falls verwendet, mit einer bestimmten Nukleinsäuresonde korrelativ ist und durch nicht-flu-

oreszente Spektrometrie oder Potentiometrie nachweisbar ist. Die DNA-Fragmente werden in vier getrennten Reaktionen reagiert, die jede einen verschieden markierten Didesoxyterminator enthalten, wobei der Marker mit einem bestimmten Didesoxynukleotid korrelativ ist und durch nicht-fluoreszente Spektrometrie oder Potentiometrie nachweisbar ist. Die DNA-Fragmente werden gemäß der Größe durch zum Beispiel Gelelektrophorese (z. B. Polyacrylamidgelelektrophorese) oder, bevorzugterweise, HPLC aufgetrennt. Die Marker werden von den aufgetrennten Fragmenten aufgespalten und durch die jeweilige Nachweisttechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen. Die nachgewiesenen Marker können mit den bestimmten zu analysierenden DNA-Fragmenten sowie der Identität des Mutanten-Nukleotids korreliert werden.

#### 10. Amplifiziertes Längenpolymorphismus (AFLP)

[0306] AFLP wurde als ein hoch sensitives Verfahren für DNA-Fingerprinting entwickelt, das in einer Vielzahl von Gebieten verwendet werden soll, einschließlich Pflanzen- und Tierzucht, medizinische Diagnose, forensische Analyse und mikrobieller Typisierung, nur um ein paar zu nennen (Vos et al., *Nucleic Acids Res.* 23: 4407–4414, 1995). Die Stärke von AFLP basiert auf den molekulargenetischen Variationen, die zwischen eng verwandten Spezies, Sorten oder Kulturen bestehen. Diese Variationen in der DNA-Sequenz werden durch die genetische Fingerprint-Technologie ausgenutzt, so daß "Fingerprints" von bestimmten Genotypen routinemäßig erzeugt werden können. Diese "Fingerprints" sind einfache RFLPs, die durch selektive PCR-Amplifikation von DNA-Restriktionsfragmenten visualisiert werden. Kurz gesagt, besteht die genetische Fingerprinting-Technologie aus den folgenden Schritten: genomische DNA wird bis zur Vollständigkeit durch zwei verschiedene Restriktionsenzyme verdaut. Spezifische Doppelstrang-Oligonukleotid-Adaptoren (~25–30 bp) werden an die verdauten DNA-Fragmente ligiert. Zu den Adaptoren homologe Oligonukleotidprimer, die jedoch Verlängerungen an dem 3'-Ende aufweisen, werden dazu verwendet, um ein Untersetz der DNA-Fragmente zu amplifizieren (ein Präamplifikationsschritt kann auch durchgeführt werden, wenn diese Verlängerung nur eine Länge von 1 bp aufweist. Die Amplifikation mit dem Primer, der eine 3-Basenpaar Verlängerung aufweist, würde folgen.) Diese Verlängerungen können in ihrer Länge von einem bis drei Basenpaare variieren, sind jedoch für einen bestimmten Primer von definierter Länge. Die Sequenz der Verlängerung kann auch von einem Primer zu einem anderen variieren, ist jedoch für einen bestimmten Primer von definierter Länge. Die Sequenz der Verlängerung kann auch von einem Primer zum anderen variieren, ist jedoch innerhalb eines Primers von einer einzelnen, definierten Sequenz. Die selektive Natur der AFLP-PCR basiert auf den 3'-Verlängerungen an den Oligonukleotid-Primern. Da diese Verlängerungen nicht homolog zu der Adaptersequenz sind, werden nur DNA-Fragmente amplifiziert werden, die komplementär zu den Verlängerungen sind, aufgrund der Unfähigkeit der Taq-DNA-Polymerase, anders als einige andere DNA-Polymerasen, DNAs zu verlängern, falls eine Fehlpaarung am 3'-Ende eines Moleküls auftritt, das synthetisiert wird. Daher wird in jeder Reaktion nur eine Untergruppe des gesamten Genoms amplifiziert. Zum Beispiel, falls zwei Basenpaar (bp)-Verlängerungen verwendet werden, wird nur eines in 256 Molekülen amplifiziert. Um die Zahl von Fragmenten weiter zu begrenzen, die wirklich visualisiert werden (so daß eine handhabbare Zahl beobachtet wird), wird nur eine dieser Primer markiert. Zuletzt werden die amplifizierten DNAs auf einem Polyacrylamidgel (Sequenziertyp) aufgetrennt und ein Autoradiogramm oder Phosphorimage wird erzeugt.

[0307] Innerhalb einer Ausführungsform der Erfindung werden Verfahren zur Bestimmung der Identität eines Nukleinsäuremoleküls zur Verfügung gestellt oder zum Nachweis eines ausgewählten Nukleinsäuremoleküls in, z. B. einer biologischen Probe, wobei die Technik des genetischen Fingerprinting verwendet wird. Kurz gesagt, umfassen solche Verfahren im allgemeinen die Schritte von Verdauen (z. B. genomischer DNA) bis zur Vollständigkeit durch zwei verschiedene Restriktionsenzyme. Spezifische Doppelstrang-Oligonukleotid-Adapter (~25–30 bp) werden an die verdauten DNA-Fragmente ligiert. Gegebenenfalls kann ein Präamplifikation unter der Verwendung von Primern mit einer 1 bp Verlängerung durchgeführt werden. Das PCR-Produkt wird dann verdünnt und markierte Primer, die homolog zu den Adaptern sind, jedoch Verlängerungen an dem 3'-Ende aufweisen, werden dazu verwendet, durch PCR eine Untergruppe der DNA-Fragmente zu amplifizieren. Die sich ergebenden PCR-Produkte werden dann in ihrer Größe nach aufgetrennt. Der Größentrennschritt kann dann durch eine Vielzahl von Verfahren durchgeführt werden, einschließlich, z. B. HPLC. Die Marker werden dann von den aufgetrennten Fragmenten abgespalten und durch die jeweilige Nachweisttechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen.

#### 11. Genexpressionsanalyse

[0308] Eine der hier beschriebenen Erfindung ist ein Hochdurchsatzverfahren zur Messung der Expression von zahlreichen Genen (1–2.000) in einer einzigen Messung. Das Verfahren hat auch die Fähigkeit, parallel durchgeführt zu werden mit mehr als 100 Proben pro Verfahren. Das Verfahren ist für Wirkstoffscreening, Entwick-

lungsbiologie, molekulare medizinische Studien und ähnliches anwendbar. Innerhalb eines Aspekts der Erfindung werden Verfahren zur Verfügung gestellt zur Analyse der Muster von Genexpression aus einer ausgewählten biologischen Probe, umfassend die Schritte von (a) Exponieren von Nukleinsäuren aus einer biologischen Probe, (b) Kombinieren der exponierten Nukleinsäuren mit einer oder mehrerer ausgewählten markierten Nukleinsonden unter Bedingungen und für eine Zeitdauer, die ausreichend ist, für die Sonden, an den Nukleinsäuren zu hybridisieren, wobei der Marker mit einer bestimmten Nukleinsäuresonde korrelativ ist und nachweisbar durch nicht-fluoreszente Spektrometrie oder Potentiometrie, (c) Abtrennen von hybridisierten Sonden von nicht-hybridisierten Sonden, (d) Abspalten des Markers von dem markiertem Fragment, (e) Nachweisen des Markers durch nicht-fluoreszente Spektrometrie oder Potentiometrie und Nachweisen des Muster der Genexpression der biologischen Probe daraus.

[0309] Innerhalb einer besonders bevorzugten Ausführungsform der Erfindung werden Tests oder Verfahren zur Verfügung gestellt, die wie folgt beschrieben werden: RNA aus einer Zielquelle wird an einen festen Träger durch einen spezifischen Hybridisierungsschritt gebunden (d. h. Fangen von poly(A) mRNA durch eine gebundene Oligo(dT)-Fangsonde). Der feste Träger wird dann gewaschen und die cDNA wird auf dem festen Träger unter der Verwendung von Standardverfahren (d. h. reverser Transkriptase) synthetisiert. Der RNA-Strang wird dann über Hydrolyse entfernt. Das Ergebnis ist die Erzeugung einer DNA-Population, die kovalent an den festen Träger immobilisiert ist, die die Diversitäts-Häufigkeit und -komplexität der RNA widerspiegelt, aus der die cDNA synthetisiert wurde. Der feste Träger wird dann mit 1 bis mehrere tausend Sonden untersucht (hybridisiert), die komplementär zu einer Gensequenz von Interesse sind. Jeder Sondentyp ist mit einem abspaltbaren Massenspektrometriemarker oder anderem Typ von abspaltbarem Marker markiert. Nach dem Untersuchungsschritt wird überschüssige oder nicht hybridisierte Sonde gewaschen, der feste Träger wird, zum Beispiel, in die Vertiefung einer Mikrotiterplatte platziert und der Massenspektrometriemarker von dem festen Träger abgespalten. Der feste Träger wird aus der Vertiefung des Probenbehälters entfernt und die Inhalte der Vertiefung werden mit einem Massenspektrometer gemessen. Das Auftreten von spezifischen Massenspektrometermarkern zeigt die Anwesenheit von RNA in der Probe an und weist nach, daß ein spezifisches Gen in einer bestimmten biologischen Probe exprimiert wird. Das Verfahren kann auch quantifizierbar sein.

[0310] Die Zusammensetzung und Verfahren zur schnellen Messung von Genexpression und der Verwendung von spaltbaren Markern können genau wie folgt beschrieben werden. Kurz gesagt, kann Gewebe (Leber, Muskel, usw.) primäre oder transformierte Zelllinien, isolierte oder gereinigte Zelltypen oder jede andere Quelle von biologischem Material, in der die Bestimmung von genetischer Expression brauchbar ist, als eine Quelle von RNA verwendet werden. In den bevorzugten Verfahren wird das biologische Quellenmaterial in der Anwesenheit eines Chaotrops lysiert, um Nukleasen und Proteasen zu unterdrücken und die stringente Hybridisierung von Zielnukleinsäure an den festen Träger zu unterstützen. Gewebe, Zellen und biologische Quellen können effektiv in 1 bis 6 molar chaotropischen Salzen (Guanidinhydrochlorid, Guanidinthocyanat, Natriumperchlorat, usw.) lysiert werden. Nachdem die quellen-biologische Probe lysiert ist, wird die Lösung mit einem festen Träger gemischt, um ein Fangen von Zielnukleinsäure zu bewirken, die in dem Lysat vorhanden ist. In einer Permutation des Verfahrens wird RNA unter der Verwendung einer gebundenen Oligo(dT)-Fangprobe gefangen. Feste Träger können Nylonperlen, Polystyrolmikroperlen, Glasperlen und Glasoberflächen oder jeden anderen Typ von festem Träger einschließen, an den Oligonukleotide kovalent angebracht werden können. Die festen Träger sind bevorzugterweise mit einem Amin-Polymer, wie zum Beispiel Polyethylen(imin), Acrylamid, Amin-Dendrimeren, usw. beschichtet. Die Amine der Polymere werden dazu verwendet, um kovalent Oligonukleotide zu immobilisieren. Die Oligonukleotide werden bevorzugterweise mit einem 5'-Amin synthetisiert (im allgemeinen ein Hexylamin, das einen 6-Kohlenstoff-Spacer-Arm und ein distales Amin einschließt). Die Oligonukleotide können 15 bis 50 Nukleotide in der Länge betragen. Die Oligonukleotide werden durch homo-bifunktionelle oder hetero-bifunktionelle kreuz-vernetzende Reagenzien, wie zum Beispiel Cyanurchlorid, aktiviert. Die aktivierten Oligonukleotide werden dann von einem überschüssigen kreuzvernetzenden Reagenz (z. B. Cyanurchlorid) durch Ausschlußchromatographie gereinigt. Das aktivierte Oligonukleotid wird dann mit den festen Trägern gemischt, um eine kovalente Anbindung zu bewirken. Nach der kovalenten Anbringung der Oligonukleotide werden die nicht reagierten Amine des festen Trägers maskiert (d. h. mit Succinianhydrid), um die positive Ladung auf dem festen Träger zu eliminieren. Die festen Träger können parallel verwendet werden und sind bevorzugterweise in einem 96-Vertiefungen- oder 384-Vertiefungen-Format konfiguriert. Die festen Träger können an Zapfen, Stämmchen oder Stäben in einer 96-Vertiefungen- oder 384-Vertiefungen-Konfiguration angebracht werden, wobei die festen Träger entweder ablösbar oder, alternativ, integral mit der besonderen Konfiguration sind. Die besondere Konfiguration des festen Trägers ist für die Funktion des Assays nicht von kritischer Wichtigkeit, sondern beeinflusst eher die Fähigkeit des Tests, an eine Automatisierung angepaßt zu werden. Die festen Träger werden mit dem Lysat für 15 Minuten bis mehrere Stunden gemischt, um ein Fangen der Zielnukleinsäure auf dem festen Träger zu bewirken. Im allgemeinen geschieht das "Fangen" der Zielnukleinsäure durch komplementäre Basenpaarung von Ziel-RNA und der auf dem festen Träger immobilisierten Fangsonde. Eine Permutation verwendet den 3'-Poly(A)-Schwanz, der an den meisten eukaryontischen Messenger-RNAs gefunden wird, um an ein immobilisiertes Oligo(dT) auf dem festen Träger zu hybridisieren.

Eine andere Permutation ist es, ein spezifisches Oligonukleotid zu verwenden oder lange Sonden (länger als 50 Basen), um eine RNA zu fangen, die eine definierte Sequenz enthält. Eine andere Möglichkeit ist es, degenerierte Primer (Oligonukleotide) zu verwenden, was das Fangen von zahlreichen verwandten Sequenzen in der Ziel-RNA-Population bewirken würde. Die Sequenzkomplexität der RNA-Population und der Typ von verwendeter Fangsonde bestimmen die Hybridisierungszeiten. Die Hybridisierungstemperaturen werden durch den Typ von verwendetem Chaotrop und der finalen Konzentration von Chaotrop diktiert (siehe Van Ness und Chen, Nuc. Acids Res. 1991, für allgemeine Anweisungen). Das Lysat wird bevorzugterweise kontinuierlich mit dem festen Träger bewegt, um eine Diffusion der Ziel-RNA zu bewirken. Sobald der Schritt des Fangens der Zielnukleinsäure vervollständigt ist, wird das Lysat vom festen Träger gewaschen und alle Chaotrope oder Hybridisierungslösung wird entfernt. Der feste Träger wird bevorzugterweise mit Lösungen gewaschen, die ionische oder nicht-ionische Detergenzien, Puffer und Salze enthalten. Der nächste Schritt ist die Synthese von DNA, die komplementär zu der gefangenen RNA ist. In diesem Schritt dient das befestigte Fangoligonukleotid als der Verlängerungsprimer für die reverse Transkriptase. Die Reaktion wird im allgemeinen bei 25–37°C durchgeführt und im allgemeinen während der Polymerisierungsreaktion geschüttelt. Nachdem die cDNA synthetisiert ist, wird sie kovalent an den festen Träger angebracht, da das Fangoligonukleotid als der Verlängerungsprimer dient. Die RNA wird dann aus der cDNA-/RNA-Duplex hydrolysiert. Der Schritt kann durch die Verwendung von Hitze durchgeführt werden, der die Duplex denaturiert oder der Verwendung von Base (d. h. 0,1 N NaOH), um die RNA chemisch zu hydrolysieren. Das Schlüsselergebnis an diesem Schritt ist es, die cDNA für die anschließende Hybridisierung mit definierten Sonden erhältlich zu machen. Der feste Träger oder das Set von festen Trägern wird dann weiter gewaschen, um RNA oder RNA-Fragmente zu entfernen. An diesem Punkt enthält der feste Träger eine ungefähre repräsentative Population von cDNA-Molekülen, die die RNA-Population im Hinblick auf Sequenzhäufigkeit, -komplexität und -diversität widerspiegeln. Der nächste Schritt ist es, ausgewählte Sonden an den festen Träger zu hybridisieren, um die Anwesenheit oder Abwesenheit und die relative Häufigkeit von spezifischen cDNA-Sequenzen zu identifizieren. Die Sonden sind bevorzugterweise Oligonukleotide mit einer Länge von 15–50 Nukleotiden. Die Sequenz der Sonden wird durch die Endverwender des Assays bestimmt. Zum Beispiel, falls der Endverwender beabsichtigt, eine Genexpression in einer entzündlichen Antwort in einem Gewebe zu untersuchen, würden die Sonden komplementär zu zahlreichen Cytokin mRNAs, RNAs, die für Enzyme kodieren, die Lipide modulieren, RNAs, die für Faktoren kodieren, die Zellen regulieren, die an einer entzündlichen Antwort beteiligt sind, usw., ausgewählt werden. Sobald ein Set von definierten Sequenzen für die Untersuchung definiert ist, wird jede Sequenz in einer Oligonukleotidsonde hergestellt und jeder Sonde ein spezifischer spaltbarer Marker zugeordnet. Der Marker/die Marker werden dann an die jeweiligen Oligonukleotid(e) angebracht. Das Oligonukleotid/die Oligonukleotide werden an die cDNA auf dem festen Träger unter geeigneten Hybridisierungsbedingungen hybridisiert. Nach Vervollständigung des Hybridisierungsschrittes wird der feste Träger gewaschen, um jede nicht-hybridisierte Sonde zu entfernen. Der feste Träger oder das Array von Trägern wird dann erhitzt, um die kovalente Bindung zwischen der cDNA und dem festen Träger zu spalten. Die markierten cDNA-Fragmente werden dann gemäß ihrer Größe durch Gelelektrophorese getrennt (z. B. Polyacrylamidgelelektrophorese) oder bevorzugterweise HPLC. Die Marker werden dann von den DNA-Molekülsonden abgespalten und durch die jeweilige Nachweistechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrometrie) nachgewiesen. Jeder vorhandene Marker wird identifiziert und die Anwesenheit (und Häufigkeit) oder Abwesenheit einer exprimierten mRNA wird nachgewiesen.

[0311] Ein alternatives Verfahren würde die markierten DNA-Sonden direkt an die befestigten mRNA-Zielmoleküle unter den geeigneten Hybridisierungsbedingungen hybridisieren. Nach Vervollständigung des Hybridisierungsschrittes wird der feste Träger gewaschen, um jede nicht hybridisierte Probe zu entfernen. Die RNA wird dann aus der DNA-Sonden/RNA-Duplex hydrolysiert. Der Schritt kann durch die Verwendung von Hitze erfolgen, die die Duplex denaturiert oder die Verwendung von Base (d. h. 0,1 N NaOH), um die RNA chemisch zu hydrolysieren. Dieser Schritt wird freie mRNA und ihre entsprechenden DNA-Sonden zurücklassen, die dann durch einen Größenabtrennschritt isoliert werden können, der im allgemeinen aus Gelelektrophorese besteht (z. B. Polyacrylamidgelelektrophorese) oder bevorzugterweise HPLC. Die Marker werden dann von den DNA-Molekülsonden abgespalten und durch die jeweilige Nachweistechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen. Jeder vorhandene Marker wird identifiziert und die Anwesenheit (und Häufigkeit) oder Abwesenheit einer exprimierten mRNA wird bestimmt.

[0312] Ein bevorzugter Genexpressionstests der vorliegenden Erfindung verwendet markierte Oligonukleotide in Verbindung mit PCR (oder anderen gleichfalls effektiven Techniken), Lambda-Exonuklease, Ultrafiltration und einem internen Standard um quantitative Information über die Genexpression zu erhalten. Dieser bevorzugte Genexpressionsassay und Unterverfahren davon werden als nächstes beschrieben.

## a. BESCHREIBUNG DES INTERNEN STANDARDVERFAHRENS

[0313] Die PCR ist ein hoch sensitives Verfahren zum Nachweis von kleinen Mengen von DNA oder RNA (durch RT-PCR). Jedoch ist die genaue und präzise Quantifizierung des Ziels schwierig, da die Menge von amplifiziertem Produkt nicht immer proportional zu der Menge von Templat ist. Dies liegt daran, daß PCR eine "Plateau-Phase" erreicht, in der fast die gesamte Menge von amplifiziertem Produkt erhalten wird, unabhängig von der Menge von Templat, nach einer bestimmten Zahl von Zyklen. Die Verwendung eines internen Standards während der PCR hilft dabei, dieses Problem durch Co-Amplifikation von bekannten Mengen eines internen Standardtemplats zu überwinden, wobei das selbe Primerset wie das Zieltemplat verwendet wird. Da beide Template das selbe Set von Primern verwenden, spiegelt das Verhältnis zwischen den Mengen der zwei amplifizierten Produkte das anfängliche Verhältnis zwischen der Menge von Ziel- und dem internen Standardtemplat vor der PCR-Amplifikation wider. Die Menge von Zieltemplat kann dann aus der bekannten Menge des internen Standardtemplats berechnet werden. Das bevorzugte interne Standardtemplat für einen PCR-Test wird sich mit einer identischen Effizienz im Vergleich zu dem Zieltemplat amplifizieren lassen. Identische Primerstellen werden in das interne Standardtemplat eingebaut, um eine Co-Amplifikation sicherzustellen. Die Region zwischen den Primerstellen und dem internen Standardtemplat sind typischerweise von dem Zieltemplat verändert (z. B. Deletion oder Addition von 10–20 Basenpaaren) um die Template durch Gelelektrophorese oder durch Restriktionsenzyme unterscheidbar zu machen. Modifikationen dieses Typs verursachen jedoch Unterschiede in der Amplifikationseffizienz der Template. Typischerweise wird eine Zahl von internen Standardtemplaten hergestellt und getestet, bis eine nahezu identische Amplifikation gefunden wird.

[0314] Ein internes Standardassayformat wird unter der Verwendung der CMST-Marker und Hybotropen entwickelt. Hybotrope sind beschrieben in, z. B., U.S.-Anmeldung Nrn. 60/026,621 (angemeldet 24. September 1996); 08/719,132 (angemeldet 24. September 1996); 08/933,924 (angemeldet 23. September 1997); 09/002,051 (angemeldet 31. Dezember 1997); und Internationale Publikation Nr. WO 98/13527, die alle hier in ihrer Gesamtheit mit aufgenommen sind. Grundsätzlich betrifft Hybotrope jede chemische Substanz, die die Enthalpie einer Nukleinsäureduplex um 20% oder mehr erhöhen kann, wenn mit einer Standardsalzlösung (d. h. 0,165 M NaCl) verglichen. Eine chemische Substanz zeigt hybotrope Eigenschaften, wenn als eine Lösung einer 18 bp Oligonukleotid-Duplex, die 50% G + C ist, eine helikale-coil Transition (HCT) von 15°C oder weniger aufweist. Die HCT ist der Unterschied zwischen den Temperaturen, bei denen 80% und 20% der Duplex einzelsträngig sind. Die Temperatur zur Anlagerung wird dann als diejenige Unterscheidungstemperatur gewählt, die eine Temperatur ist, bei der eine Hybridisierungsreaktion durchgeführt wird, die eine nachweisbare Unterscheidung zwischen einer fehlgepaarten Duplex und einer perfekt gepaarten Duplex ermöglicht. Ein Bereich von Temperaturen erfüllt diese Kriterien einer Unterscheidungstemperatur.

[0315] Da hoch spezifische Hybridisierung unter der Verwendung von Hybotropen durchgeführt werden kann, kann ein bevorzugtes internes Standardtemplat verwendet werden, das dieselbe Länge aufweist und nur ein Basenpaar unterschiedlich von dem Zieltemplat ist. Dies ist ein bevorzugter interner Standard, da er a priori identisch mit dem Zieltemplat co-amplifizieren wird. Ein bevorzugtes Verfahren des Nachweises dieses Typs von Standard und Zielamplikongemisch ist es, hybotrope Puffer zu verwenden, um eine spezifische Hybridisierung von Amplikon-spezifisch markierten Oligonukleotiden (die bevorzugterweise CMST-markiert sind, d. h. durch Massenspektrometrie nachweisbare Marker aufweisen), an deren jeweilige Amplikons, um die Menge von Amplikon nachzuweisen, wird es Amplikonspezifische Oligonukleotid mit einem einmaligen Massenmarker markiert. Die Amplikon-Quantifizierung wird durch Messen des Massenspektrometriesignals abgeleitet, das mit dem internen Standard und den Zielamplikons assoziiert ist. Eine bekannte Menge von internem Standard wird in eine Probe gegeben, und die Menge von Zielamplikon wird unter der Verwendung der folgenden Gleichung und der Verwendung der Daten bestimmt, die im Anschluß an die Lambda-Exonuklease und Ultrafiltrationsverfahren erhalten werden, die als nächstes beschrieben werden.

[0316]

$$\text{Ziel}_{\text{Massensignal}} / \text{interner Standard}_{\text{Massensignal}} \times \text{interner Standard}_{\text{Menge}} = \text{Ziel}_{\text{Menge}}$$

## b. BESCHREIBUNG DES LAMBDA EXONUKLEASE VERFAHRENS

[0317] Nach Inkubation mit einer DNA-Duplex verdaut Lambda-Exonuklease selektiv einen Strang von einem 5'-phosphoryliertem Ende her, was ein für DNA-Sequenzierung geeignetes einzelsträngiges Templat zurückläßt. Die Lambda-Exonuklease stellt eine einzelsträngige Sequenzierungstemplate her, ohne die Mühe von herkömmlichen biologischen Verfahren oder der Lästigkeit der Optimierung von asymmetrischer PCR. Durch Amplifikation von DNA in der Anwesenheit eines Primers, der ein 5'-terminales Phosphat enthält, kann ein Fachmann einer DNA-Duplex mit einem 5'-phosphoryliertem Ende erzeugen. Danach wird das PCR-Produkt entweder durch Fällung oder Gelfiltration gereinigt, um restliche Primer und andere Reaktionskomponenten zu entfernen. Der phosphorylierte Strang der DNA-Duplex wird dann selektiv mit Lambda-Exonuklease verdaut,

was ein einzelsträngiges nicht-phosphoryliertes Templat zurückläßt, das für die Sequenzierung geeignet ist. Nach der Hitzeinaktivierung der Lambda-Endonuklease kann die konzentrierte, einzelsträngige DNA direkt zu einer Hybridierungsreaktion hinzugefügt werden, die herkömmliche Techniken anwendet.

### c. BESCHREIBUNG DES ULTRAFILTRATIONSVERFAHRENS

[0318] Genomische DNA, wie aus weißen Blutzellen geerntet, kann gemäß den Standardverfahren von RNase-Behandlung, Proteinase-K-Verdau und Phenol : Chloroform-Extraktion, gefolgt durch Fällung mit Natriumacetat und Ethanol, gereinigt werden. Die Konzentrationen werden durch Spektrophotometrie bestimmt und Arbeitsverdünnungen von 0,01 µg/µl werden hergestellt. 50 µl der DNA-Proben werden in 96-Vertiefungs-"Mutter"-Platten ausgebracht. "Tochter"-Platten, die für die Amplifikationsreaktion verwendet werden, werden durch transferieren von 1,5 µl von DNA aus der Mutterplatte aus den Vertiefungen von entweder einer 96-Vertiefungsplatte oder 192-Vertiefungsplatte in einer identischen Konfiguration unter der Verwendung von 8-Kanal-Pipetten präpariert. Eine 18 µl Lage von flüssigem Wachs (MJ Research) wird dann hinzugefügt und dann wird diese Platte bei 4 Grad gelagert, wobei sich das Wachs verfestigt, was eine Verdampfung verhindert.

[0319] Um die PCR anzusetzen, werden die Tochterplatten entfernt und auf Eis plaziert, um das Wachs festzuhalten. Dieses bildet eine Barriere zwischen der Templat-DNA und anderen Komponenten der Reaktion, bis die Platte in dem Thermal-Cycler plaziert wird und die Reaktion durch Erhitzen während des ersten Zyklus initiiert wird. Die PCR wird in einer Gesamtzahl von 10–50 µl pro Reaktion durchgeführt. Mastermixlösungen werden vorher in Vorratslösungen hergestellt und in Röhrchen aufgeteilt, die alle der Komponenten der PCR, außer den Marker-spezifischen Primern enthalten. Die Hauptzahl der Reaktionen werden mit M13-verlängerten Primern durchgeführt. M13-verlängerte Primer sind eine Modifikation des Standard-PCR-Primerpaars. Die Modifikation ist die Hinzufügung von 17 Nukleotiden an das 5'-Ende des Vorwärtsprimers. Die 17-Nukleotidsequenz ist komplementär zu dem m13-Sequenzierungsprimer und besitzt die Sequenz: 5'-(NH<sub>2</sub>)-AGG GTT TT-CCCA GTC ACG AC-3'. Die Modifikation erlaubt die Verwendung eines dritten Oligonukleotidprimers in einer PCR-Reaktion. Der dritte Primer ist typischerweise gemäß den hier beschriebenen Verfahren markiert.

[0320] Ultrafiltration ist das herkömmliche Verfahren zur Konzentration und Entsalzung von Proteinen. Es ist auch ein effiziente Alternative zur Ethanol-fällung von Nukleinsäuren, insbesondere für kleine Mengen und besonders, wenn immer Nukleinsäuren nur gefällt werden, um Lösungsmittel auszuwechseln. Für Proben, die Phosphat oder 10 mM EDTA enthalten, kann die Ultrafiltration eine bemerkenswert zeitsparende Methodologie sein. Herkömmlich erforderten solche Proben die vorhergehende Dialyse, um die Co-Fällung von Salzen mit Nukleinsäuren während der Ethanol-fällung zu vermeiden. Zentrifugal-Mikrokonzentratoren entsalzen und konzentrieren Oligonukleotidsonden oder einzelsträngige Amplikons (und Nukleinsäuren) in einem einfachen Schritt.

[0321] Microcon® Mikrokonzentratoren sind ideal zur Konzentration von 50–500 µl Proben. Bei der zentrifugalen Ultrafiltration wird die DNA durch die Membran zurückgehalten. Lösungsmittel und Salze passieren und werden entfernt. Eine zweite umgedrehte Zentrifugation von Microcon stellt einen maximalen DNA-Rückerhalt der Sonde oder Nukleinsäure von Interesse sicher. Bei der Konzentration der Oligomerproben ist es wichtig, hohe Salzkonzentrationen zu vermeiden, die eine Bindung von einzelsträngigen Nukleinsäure an Zellulose-basierte Ultrafiltrationsmembranen unterstützen.

[0322] Typische Ultrafiltrationsbedingungen sind wie folgt: eine DNA-Lösung (500 µl) wird einem Microcon-30-Konzentrator für 10 Minuten bei 12.000 × g zentrifugiert; eine 500 µl Lösung von Oligonukleotiden in TE-Puffer wird in einer Microcon-3-Einheit für 45 Minuten bei 12.000 × g zentrifugiert. Die Überstände können durch Umdrehen der Vorrichtungen und Zentrifugieren bei 500–1.000 × g für 2 Minuten zurückerhalten werden. Im Gegensatz dazu erfordert Gelelektrophorese Fraktionierung, Elution und Entsalzung des Fragments aus der Gelscheibe. Sie erfordert auch genügend Material, um zu visualisieren, was manchmal schwierig zu erhalten ist, insbesondere im Fall von cDNA. Beide Verfahren (Gelelektrophorese und Dialyse) sind zeitaufwendig und schließen viele Probenverarbeitungsschritte ein.

[0323] In einer typischen Hybridisierungsreaktion werden die Sonden in 50–100-fachen molaren Überschoß von DNA-Fragmentkonzentration hinzugefügt. Aus diesem Grund ist es erforderlich, die überschüssige nicht-hybridisierte Sonde zu entfernen. Herkömmliche Verfahren zur Sondenentfernung schließen Gelfiltration-chromatographie oder Gelelektrophorese ein. Die Gelfiltration erfordert, daß viele Fraktionen gesammelt, analysiert, vereinigt und gefällt werden, was für Hochdurchsatzassays nicht anwendbar ist. Jedoch ist die Ultrafiltration eine effektive Alternative für ein schnelles Entfernen von überschüssiger Sonde von PCR-Primern. Bei Amicon's Centricon wegwerfbarem Konzentrator, würde das Reaktionsgemisch durch ein Ultrafiltrationsmembran filtriert, was zu der Entfernung von Puffer und nicht-hybridisierter Sonde oder nicht-verlängerten Primern führt. Die konzentrierten Fragmente werden durch die Membran zurückgehalten. Antreibende Kraft für die Filtration wird durch die Zentrifugation in einem Rotor mit fixiertem Winkel bei 1.000-5.000 g zur Verfügung gestellt. Herkömmliche Verfahren erfordern typischerweise 24 Stunden für die Probenverarbeitung und 2–3 Stunden Arbeitszeit. Bei Centricon benötigen die Proben typischerweise weniger als ein paar Stunden, um die-

se zu reinigen. Die Probenhandhabung ist minimal und viele Proben können zur selben Zeit verarbeitet werden. Als Referenz siehe Krowczynska, A. M., "Efficient Purification of PCR Products using Ultrafiltration" *Bio-Techniques* 13(2): 286–289, 1992.

## 12. Hybridisierungstechniken

[0324] Die erfolgreiche Klonierung und Sequenzierung eines Gens führt zu der Untersuchung seiner Struktur und der Expression dadurch, daß es möglich gemacht wird, das Gen nachzuweisen oder seine mRNA in einem großen Pool von nicht-verwandten DNA- oder RNA-Molekülen. Die Menge von mRNA, die für ein spezifisches Protein in Gewebe kodiert, ist ein wichtiger Parameter für die Aktivität eines Gens und kann signifikant mit der Aktivität von funktionierenden Systemen zusammenhängen. Seine Regulation ist abhängig von der Interaktion von Sequenzen innerhalb des Gens (cis-wirkende Elemente) und Sequenz-spezifischen DNA-bindenden Proteinen (trans-wirkenden Faktoren), die Gewebe-spezifisch aktiviert werden, oder durch Hormone und Second Messenger Systeme.

[0325] Verschiedene Techniken zur Analyse eines bestimmten Gens, seine regulatorischen Sequenzen, seiner spezifischen mRNA und der Regulation seiner Expression sind erhältlich, diese schließen Southern- oder Northern-Blot-Analyse und Ribonuklease-(Rnase)-Protection-Assay ein.

[0326] Veränderungen in der Nukleotidzusammensetzung eines bestimmten Gens können von großer pathophysiologischer Relevanz sein. Wenn in den nicht-kodierenden Regionen (5'-, 3'-flankierenden Regionen und Intron) lokalisiert, können sie die Regulation der Genexpression beeinflussen, was abnormale Aktivierung oder Inhibierung verursacht. Wenn in den kodierenden Regionen des Gens (Exone) lokalisiert, können sie zu Veränderung der Proteinfunktion oder disfunktionellen Proteinen führen. Daher kann eine bestimmte Sequenz innerhalb eines Gens mit einer spezifischen Erkrankung korrelieren und kann als ein Marker der Erkrankung brauchbar sein. Ein primäres Ziel der Forschung im medizinischen Bereich ist es daher, diejenigen genetischen Variationen als diagnostische Tools nachzuweisen und wichtige Informationen für das Verständnis von pathophysiologischen Phänomenen zu erhalten. Das Basisverfahren für die Analyse einer Population im Hinblick auf die Variationen innerhalb eines bestimmten Gens ist DNA-Analyse unter der Verwendung der Southern-Blot-Technik. Kurz gesagt, wird präparierte genomische DNA mit einem Restriktionsenzym (RE) verdaut, was zu einer großen Zahl von DNA-Fragmenten von verschiedenen Längen führt, die durch die Anwesenheit der spezifischen Erkennungsstelle des RE im Genom bestimmt werden. Allele eines bestimmten Gens mit Mutationen innerhalb dieser Restriktionsstelle werden in Fragmente verschiedener Zahl und Länge gespalten werden. Dies wird Restriktionslängenpolymorphismus (RFLP) genannt und kann ein wichtiger diagnostischer Marker mit vielen Anwendungen sein. Das zu analysierende Fragment muß aus dem Pool von DNA-Fragmenten abgetrennt werden und von anderen DNA-Spezies unter der Verwendung einer spezifischen Sonde unterschieden werden. Daher wird die DNA einer elektrophoretischen Fraktionierung unter Verwendung eines Agarosegels unterzogen, gefolgt von Transfer und Fixierung an eine Nylon- oder Nitrozellulosemembran. Die fixierte, Einzelsträngige DNA wird an eine markierte DNA hybridisiert, die komplementär zu der nachzuweisenden DNA ist. Nach Entfernen von nicht-spezifischen Hybridisierungen kann das DNA-Fragment von Interesse gemäß den Sondencharakteristika (Autoradiographie oder Phosphorimageanalyse) visualisiert werden.

[0327] Innerhalb einer Ausführungsform der Erfindung werden Verfahren zur Bestimmung der Identität eines Nukleinsäuremoleküls oder zum Nachweis eines ausgewählten Nukleinsäuremolekül in, zum Beispiel, einer biologischen Probe zur Verfügung gestellt, wobei die Techniken ähnlich zu Southern-Blotting verwendet werden. Kurz gesagt umfassen solche Verfahren im allgemeinen die Schritte von Erzeugen einer Reihe von markierten Nukleinsäurefragmenten, in denen die erzeugten Fragmente mit Restriktionsenzymen verdaut werden. Die markierten Fragmente werden durch Durchführen eines Hybridisierungsschritts der markierten Sonden mit der verdauten Zielnukleinsäure erzeugt. Der Hybridisierungsschritt kann vor oder nach dem Restriktionsnukleaseverdau erfolgen. Die sich ergebenden verdauten Nukleinsäurefragmente werden dann nach ihrer Größe getrennt. Der Größentrennungsschritt kann dann durch zum Beispiel Gelelektrophorese (z. B. Polyacrylamid-gelelektrophorese) oder, bevorzugterweise HPLC, erfolgen. Die Marker werden dann von den abgetrennten Fragmenten abgespalten und dann werden die Marker durch die jeweilige Nachweisthechnologie nachgewiesen (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie).

[0328] Die Anwesenheit und Quantifizierung eines spezifischen Gentranskripts und seine Regulierung durch physiologische Parameter können mittels Northern-Blot und RNase-Protection-Assay analysiert werden. Die prinzipielle Basis dieser Verfahren ist die Hybridisierung eines Pools von gesamter zellulären DNA an eine spezifische Sonde. Bei der Northern-Blot-Technik wird Gesamt-RNA eines Gewebes unter der Verwendung eines HPLC- oder LC-Verfahrens fraktioniert, an eine markierte Antisense-RNA (cRNA) hybridisiert, die komplementär zu der nachzuweisenden RNA ist. Durch Anwendung von stringenten Waschbedingungen werden nicht-spezifisch gebundene Moleküle eliminiert. Spezifisch gebundene Moleküle würden anschließend gemäß

des Typs von verwendeter Sonde (Massenspektrometrie oder mit einem elektrochemischen Detektor) nachgewiesen. Zusätzlich kann die Spezifität durch Vergleichen der Größe der nachgewiesenen mRNA und der vorhergesagten Länge der mRNA von Interesse kontrolliert werden.

[0329] Innerhalb einer Ausführungsform der Erfindung werden Verfahren zum Nachweis der Identität eines Ribonukleinsäuremoleküls oder zum Nachweis eines ausgewählten Ribonukleinsäuremoleküls in, zum Beispiel, einer biologischen Probe zur Verfügung gestellt, wobei die Techniken ähnlich zu Northern-Blotting verwendet werden. Kurz gesagt umfassen solche Verfahren im allgemeinen die Schritte von Erzeugen einer Reihe von markierten RNA-Molekülen durch Durchführen eines Hybridisierungsschrittes der markierten Sonden mit der Ziel-RNA. Die Ziel-RNA-Moleküle werden dann nach ihrer Größe aufgetrennt. Der Größenaufreinigungsschritt kann dann, zum Beispiel, bevorzugterweise, HPLC erreicht werden. Die Marker werden dann von den getrennten RNA-Molekülen abgespalten und dann werden die Marker durch die jeweilige Nachweistechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen.

[0330] Das meisten spezifische Verfahren zum Nachweis einer mRNA-Spezies ist der RNase-Protection-Assay. Kurz gesagt, wird Gesamt-RNA aus einem Gewebe oder einer Zellkultur an eine markierte spezifische mRNA von vollständiger Homologie hybridisiert. Die Spezifität wird durch anschließenden RNase-Verdau erreicht. Nicht-hybridisierte Einzel-strängige RNAs und nicht-spezifisch hybridisierte Fragmente mit sogar kleinen Fehlpaarungen werden erkannt und gespalten, während doppelsträngige RNA von vollständiger Homologie nicht für das Enzym zugänglich ist, und geschützt werden wird. Nach Entfernen der RNase durch Proteinase-K-Verdau und Phenol-Extraktion kann das spezifisch geschützte Fragment von den Abbauprodukten abgetrennt werden, gewöhnlicherweise durch HPLC.

[0331] Innerhalb einer Ausführungsform der Erfindung werden Verfahren zur Bestimmung der Identität eines Ribonukleinsäuremoleküls oder zum Nachweis eines ausgewählten Ribonukleinsäuremolekül in, zum Beispiel, einer biologischen Probe zur Verfügung gestellt, wobei die Technik des RNase-Protection-Assays verwendet wird. Kurz gesagt umfassen solche Verfahren im allgemeinen die Schritte von Gesamt-RNA aus einem Gewebe oder einer Zellkultur, die an eine markierte spezifische cRNA von vollständiger Homologie hybridisiert wird, einen RNase-Verdau, Behandlung mit Proteinase K und eine Phenol-Extraktion. Das markierte geschützte RNA-Fragment wird aus den Abbauprodukten isoliert. Der Größenabtrennungsschritt kann zum Beispiel durch LC oder HPLC erreicht werden. Der Marker wird dann von dem abgetrennten RNA-Molekül abgespalten und wird dann durch die jeweilige Nachweistechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen.

### 13. Mutationsnachweistechniken

[0332] Der Nachweis von Erkrankungen ist bei der Vorbeugung und Behandlungen zunehmend von Wichtigkeit. Während es für multifaktorielle Erkrankungen schwierig ist genetische Tests dafür zu entwickeln, werden mehr als 200 bekannte menschliche Erkrankungen durch einen Defekt in einem einzelnen Gen, oft eine Veränderung eines einzelnen Aminosäurerests (Olsen, *Biotechnology: An industry comes of age*, National Academic Press, 1996) verursacht. Viele dieser Mutationen führen zu einer veränderten Aminosäure, die einen Erkrankungszustand verursacht.

[0333] Sensitive Mutationsnachweistechniken bieten außerordentliche Möglichkeiten für Mutationsscreening. Zum Beispiel, können Analysen durchgeführt werden sogar vor der Implantation eines befruchteten Eis (Holding und Monk, *Lancet* 3: 532, 1988). Erhöhte effiziente genetische Tests können auch ein Screening auf onkogene Mutationen in Zellen ermöglichen, die vom respiratorischen Trakt oder der Blase abgestoßen wurden, in Verbindung mit Vorsorgeuntersuchungen (Sidransky et al., *Science* 252: 706). Auch sind, wenn ein unbekanntes Gen eine genetische Erkrankung verursacht, Verfahren zur Überwachung von DNA-Sequenzvarianten brauchbar, um die Vererbung von Erkrankung durch genetische Kopplungsanalyse zu untersuchen. Jedoch verursacht der Nachweis und die Diagnose von Mutationen in einzelnen Genen technische und ökonomische Herausforderungen. Mehrere verschiedene Ansätze wurden verfolgt, jedoch ist keiner sowohl effizient und kostengünstig genug für eine wirklich weit verbreitete Anwendung.

[0334] Mutationen, die ein einzelnes Nukleotid einschließen, können in einer Probe durch physikalische, chemische oder enzymatische Mittel identifiziert werden. Im allgemeinen können Verfahren zum Mutationsnachweis in Scan-Techniken eingeteilt werden, die geeignet sind, um vorher unbekannte Mutationen zu identifizieren und Techniken, die so aufgebaut sind, um bekannte Sequenzvarianten nachzuweisen, zu unterscheiden oder zu quantifizieren. Verschiedenen Scan-Techniken für den Mutationsnachweis wurden in Heteroduplexen von fehlgepaarten komplementären DNA-Strängen, die von Wildtyp- und Mutantensequenzen abgeleitet waren und ein abnormales Verhalten, insbesondere wenn denaturiert, zeigten, entwickelt. Dieses Phänomen wird bei der denaturierenden und Temperaturgradient-Gelelektrophorese-(DGGE und TGGE)-Verfahren ausgenutzt. Duplexe, die sogar in einer einzelnen Nukleotidposition fehlgepaart sind, können teilweise denaturieren, was zu einer verzögerten Wanderung führt, wenn in einem ansteigend denaturierenden Gradienten gelelekt-



rophoriert (Myers et al., Nature 313: 495, 1985; Abrades et al., Genomics 7: 463, 1990; Henco et al., Nucl. Acids Res. 18: 6733, 1990). Obwohl Mutationen nachgewiesen werden können, wird keine Information im Hinblick der präzisen Lokalisierung einer Mutation erhalten. Die Mutanten-Formen müssen weiterhin isoliert werden und einer DNA-Sequenzanalyse unterzogen werden. Alternativ kann RNase A eine Heteroduplex einer RNA-Sonde und eines Zielstranges an einer Position spalten, wo die zwei Stränge nicht genau gepaart sind. Diese Spaltstelle kann dann durch Elektrophorese der denaturierten Probe bestimmt werden. Jedoch können einige Mutationen dem Nachweis entkommen, da nicht alle Fehlpaarungen durch RNase A effizient gespalten werden. Fehlgepaarte Basen in einer Duplex sind ebenfalls gegenüber chemischer Modifikation empfindlich. Solche Modifikationen können die Stränge gegenüber Spaltung an der Stelle der Fehlpaarung empfindlich machen oder dazu führen, daß eine Polymerase in einer anschließenden Verlängerungsreaktion dort stoppt. Die chemische Spalttechnik ermöglicht die Identifizierung einer Mutation in Zielsequenzen von bis zu 2 kb und sie stellt Informationen über die ungefähre Lokalisierung von fehlgepaarten Nukleotiden zur Verfügung (Cotton et al., PNAS LISA 85: 4397, 1988; Ganguly et al., Nucl. Acids Res. 18: 3933, 1991). Jedoch ist diese Technik arbeitsintensiv und kann die präzise Lokalisierung der Mutation vielleicht nicht identifizieren.

[0335] Eine alternative Strategie zum Nachweis einer Mutation in einem DNA-Strang ist durch Substituierung (während der Synthese) eines der normalen Nukleotide mit einem modifizierten Nukleotid, wobei das Molekulargewicht oder andere physikalische Parameter des Produkts verändert werden. Ein Strang mit einer erhöhten oder verringerten Zahl dieses modifizierten Nukleotids in Bezug auf die Wildtyp-Sequenz zeigt eine veränderte elektrophoretische Mobilität (Naylor et al., Lancet 337: 636, 1991). Diese Technik weist die Anwesenheit von einer Mutation nach, stellt jedoch nicht die Lokalisierung zur Verfügung.

[0336] Zwei andere Strategien visualisieren Mutationen in einem DNA-Fragment durch veränderte Gelmigration. In der Einzel-Strang-Polymorphismustechnik (SSCP), verursachen Mutationen meiner Meinung nach, daß denaturierte Stränge verschiedener Sekundärstrukturen annehmen, wodurch die Mobilität während nativer Gelelektrophorese beeinflusst wird. Heteroduplex-DNA-Moleküle, die interne Fehlpaarungen aufweisen, können auch von korrekt gepaarten Molekülen durch Elektrophorese abgetrennt werden (Orita, Genomics 5: 874, 1989; Keen, Trends Genet. 7: 5, 1991). Wie mit den oben diskutierten Techniken, kann die Anwesenheit einer Mutation bestimmt werden, jedoch nicht Lokalisierung. Genauso können viele dieser Techniken nicht zwischen einer einzelnen und multiplen Mutationen unterscheiden. Alle der oben genannten Techniken zeigen die Anwesenheit einer Mutation in einem begrenzten Segment von DNA an und einige von ihnen ermöglichen die ungefähre Lokalisierung innerhalb des Segments. Jedoch ist immer noch Sequenzanalyse erforderlich, um den Effekt der Mutation auf das Kodierungspotential des Fragments zu ermitteln. Sequenzanalyse ist sehr wirksam, was zum Beispiel ein Screening auf die selbe Mutation in anderen Individuen einer beeinträchtigten Familie ermöglicht, wodurch ein Erkrankungsfortschreiten im Falle von maligner Erkrankung beim Nachweis von restlichen malignen Zellen im Knochenmark vor autologer Transplantation verfolgt wird. Trotz dieser Vorteile ist es aufgrund der hohen Aufgaben, die enthalten sind, unwahrscheinlich, daß das Verfahren als ein Routine-diagnostisches Verfahren angenommen werden wird.

[0337] Eine große Anzahl von anderen Techniken wurden entwickelt, um bekannte Sequenzvarianten zu analysieren. Die Automatisierung und Wirtschaftlichkeit sind sehr wichtige Überlegungen für diese Typen von Analysen, die für das Screening von Individuen und der allgemeinen Population angewendet werden können. Keine der unten diskutierten Techniken kombiniert Wirtschaftlichkeit und Automatisierung mit der erforderlichen Spezifität.

[0338] Mutationen können über ihre destabilisierenden Effekte auf die Hybridisierung von kurzen Oligonukleotidsonden an eine Zielsequenz identifiziert werden (siehe Wetmur, Crit. Rev. Biochem. Mol. Biol. 26: 227, 1991). Im allgemeinen schließt diese Technik, Allel-spezifische Oligonukleotidhybridisierung, die Amplifikation von Zielsequenzen und anschließende Hybridisierung von kurzen Oligonukleotidsonden ein. Ein amplifiziertes Produkt kann daher auf viele mögliche Sequenzvarianten durch Bestimmen seines Hybridisierungsmusters an Array von immobilisierten Oligonukleotidsonden gescannt werden. Jedoch hängt das Etablieren von Bedingungen, die zwischen einer Zahl von anderen Strategien zur Nukleotidsequenz-Unterscheidung unterscheiden, alle von Enzymen ab, um Sequenzunterschiede zu identifizieren (Saiki, PNAS USA 86: 6230, 1989; Zhang, Nucl. Acids Res. 19: 3929, 1991).

[0339] Zum Beispiel erkennen Restriktionsenzyme Sequenzen von ungefähr 4–8 Nukleotiden. Basierend auf einem durchschnittlichen G+C-Gehalt können ungefähr die Hälfte der Nukleotidpositionen in einem DNA-Fragment mit einem Panel mit 100 Restriktionsenzymen überwacht werden. Als eine Alternative können künstliche Restriktionsenzyme-Erkennungssequenzen um eine variable Position herum durch Verwendung von teilweise fehlgepaarten PCR-Primern erzeugt werden. Mit dieser Technik kann entweder die Mutanten oder die Wildtyp-Sequenz allein erkannt werden und durch ein Restriktionsenzym nach Amplifikation gespalten werden (Chen et al., Anal. Biochem. 195: 51, 1991; Levi et al., Cancer Res. 51: 3497, 1991). Ein anderes Verfahren, nutzt die Fähigkeit aus, daß ein Oligonukleotidprimer, der an der 3'-vorletzten Position mit einer Zielsequenz fehlgepaart ist, eine verringerte Kapazität aufweist, als ein Primer in der PCR zu dienen. Jedoch sind einige 3'-Fehlpaarungen, bemerkenswerterweise G-T, weniger inhibitorisch als andere, was dessen Brauchbarkeit

begrenzt. In einem Ansatz um dieses Verfahren zu verbessern, werden zusätzliche Fehlpaarungen in dem Primer an der dritten Position von dem 3'-Ende eingefügt. Dieses führt zu zwei fehlgepaarten Positionen in dem 3'-Nukleotiden des Primers, der mit einer allelischen Variante hybridisiert und eine Fehlpaarung in der dritten Position von dem 3'-Ende wenn der Primer an die andere allelische Variante hybridisiert (Newton et al., Nucl. Acids Res. 17: 2503, 1989). Es ist erforderlich, Amplifikationsbedingungen zu definieren, die signifikant die Amplifikation einer 1 bp – Fehlpaarung bevorzugen.

[0340] DNA-Polymerasen wurden auch dazu verwendet, um allelische Sequenzvarianten zu unterscheiden, durch Bestimmen welches Nukleotid an den Oligonukleotidprimer angefügt wird, unmittelbar stromaufwärts einer variablen Position in dem Zielstrang.

[0341] Ein Ligationsassay wurde entwickelt. In diesem Verfahren werden zwei Oligonukleotidsonden, die in unmittelbarer Nebeneinanderstellung auf einem Zielstrang hybridisieren durch eine DNA-Ligase verbunden. Die Ligation ist inhibiert, falls eine Fehlpaarung auftritt, wo die zwei Oligonukleotidsonden aneinander grenzen.

#### 14. Assays für den Mutationsnachweis

[0342] Mutationen sind eine Einzel-Basenpaar-Veränderung in genomischer DNA. Innerhalb des Kontext dieser Erfindung werden die meisten solcher Veränderungen leicht durch Hybridisierung mit Oligonukleotiden nachgewiesen, die zu der fraglichen Sequenz komplementär sind. In dem hier beschriebenen System werden zwei Oligonukleotide verwendet um eine Mutation nachzuweisen. Ein Nukleotid besitzt die Wildtyp-Sequenz und das andere Oligonukleotid besitzt die Mutantensequenz. Wenn die zwei Oligonukleotide als ein Sonden auf einer Wildtyp-zielgenomischen Sequenz verwendet werden, wird das Wildtyp-Oligonukleotid eine perfekt Basen-gepaarte Struktur ausbilden und die Mutanten-Oligonukleotidsequenz wird eine Duplex mit einer Einzelbasepaarfehlpaarung bilden. Wie oben diskutiert – erlaubt eine 6 bis 7°C Differenz in der T<sub>m</sub> einer Wildtyp gegenüber einer fehlgepaarten Duplex die leichte Identifizierung und Unterscheidung der zwei Typen von Duplices. Um diese Unterscheidung zu bewirken, wird eine Hybridisierung bei der T<sub>m</sub> der fehlgepaarten Duplex in der jeweiligen hybotropen Lösung durchgeführt (siehe, z. B. U.S.-Anmeldung Nrn. 60/026,621 (angemeldet 24. September 1996); 08/719,132 (angemeldet 24. September 1996); 08/933,924 (angemeldet 23. September 1997); 09/002,051 (angemeldet 31. Dezember 1997) und Internationale Publikations-Nr. WO 98/13527, die alle hier in ihrer Gesamtheit für eine Beschreibung von hybotropen Lösungen aufgenommen sind). Das Ausmaß der Hybridisierung wird dann für das Set von Oligonukleotidsonden gemessen. Wenn das Verhältnis des Ausmaßes der Hybridisierung der Wildtyp-Sonde zu der fehlgepaarten Sonde gemessen wird, wird ein Wert von 10/1 bis größer als 20/1 erhalten. Diese Typen von Ergebnissen ermöglichen die Entwicklung von robusten Assays für den Mutationsnachweis.

[0343] Für beispielhafte Zwecke verwendet ein Testformat für den Mutationsnachweis eine Zielnukleinsäure (z. B. genomische DNA) und Oligonukleotidsonden, die den Bereich von Interesse überspannen. Die Oligonukleotidsonden sind größer oder gleich zu 24 Nukleotiden in ihrer Länge (mit einem Maximum von ungefähr 36 Nukleotiden) und markiert mit einem Fluorochrom an dem 3' oder 5'-Ende der Oligonukleotidsonde. Die Zielnukleinsäure wird über die Lyse von Gewebekulturzellen, Geweben, Organismen, usw., in der jeweiligen Hybridisierungslösung erhalten. Die lysierte Lösung wird dann auf eine Temperatur erhitzt, die die Zielnukleinsäure denaturiert (15–25°C oberhalb der T<sub>m</sub> der Zielnukleinsäureduplex). Die Oligonukleotidsonden werden bei der Denaturierungstemperatur hinzugefügt und die Hybridisierung wird bei der T<sub>m</sub> der fehlgepaarten Duplex für 0,5 bis 24 Stunden durchgeführt. Die genomische DNA wird dann gesammelt und durch einen GF/C (GF/B und ähnliches) Glasfaserfilter hindurch passiert. Der Filter wird dann mit der jeweiligen Hybridisierungslösung gewaschen, um alle nicht-hybridisierten Oligonukleotidsonden zu entfernen (RNA, kurze Oligos und Nukleinsäure bindet nicht an Glasfaserfilter und unter diesen Bedingungen). Die Hybridisierungsoligosonde kann dann von der Ziel-DNA termisch eluiert werden und gemessen werden (durch Fluoreszenz z. B.). Für Tests, die sehr hohe Spiegel an Sensitivität erfordern, werden die Sonden konzentriert und gemessen.

[0344] Andere hochsensitive Hybridisierungsprotokolle können verwendet werden. Die Verfahren der folgenden Erfindung ermöglichen es einem, leicht auf eine Nukleinsäure zu testen, die eine Mutation enthält, von der vermutet wird, daß sie in Zellen, Proben, usw. vorhanden ist, d. h. einer Zielnukleinsäure. Die Zielnukleinsäure enthält eine Nukleotidsequenz von Desoxyribonukleinsäure (DNA) oder Ribonukleinsäure (RNA), deren Anwesenheit von Interesse ist und deren Anwesenheit oder Abwesenheit in dem Hybridisierungsassay nachgewiesen werden soll. Die Hybridisierungsverfahren der folgenden Erfindung können auch auf ein komplexes biologisches Gemisch von Nukleinsäuren (RNA und/oder DNA) angewendet werden. Solch ein komplexes biologisches Gemisch schließt einen weiten Bereich von eukaryontischen und prokaryontischen Quellen, einschließlich Protoplasten ein, und/oder andere biologische Materialien, die Polynukleotidnukleinsäure enthalten. Das Verfahren ist daher auf Gewebekulturzellen, tierische Zellen, tierisches Gewebe, Blutzellen (z. B. Retikulozyten, Lymphozyten), Pflanzenzellen, Bakterien, Hefen, Viren, Mykoplasmen, Protozoen, Pilze und ähnliches anwendbar. Durch Nachweisen einer spezifischen Hybridisierung zwischen Nukleinsäuresonden einer bekannten Quelle kann die spezifische Anwesenheit einer Zielnukleinsäure etabliert werden. Ein typisches Hybridisie-

rungsassayprotokoll zum Nachweis einer Zielnukleinsäure in einer komplexen Population von Nukleinsäuren wird wie folgt geschrieben: Zielnukleinsäure werden in ihrer Größe auf einer LC oder HPLC aufgetrennt, kloniert und isoliert, unterteilt in Pools oder als eine komplexe Population belassen. Innerhalb einer Ausführungsform der Erfindung werden Verfahren zum Nachweis der Identität eines Nukleinsäuremoleküls oder zum Nachweis eines ausgewählten Nukleinsäuremoleküls in, z. B. einer biologischen Probe zur Verfügung gestellt, wobei die allgemeinen Techniken von Hybridisierungsassay verwendet werden. Kurz gesagt umfassen solche Verfahren im allgemeinen die Schritte von Klonierung und Isolierung von Zielnukleinsäuren, Unterteilung in Pools oder Belassen als eine komplexe Population. Die Zielnukleinsäuren werden dann mit markierten Oligonukleotidsonden unter oben beschriebenen Bedingungen hybridisiert. Die Zielnukleinsäuren werden ihrer Größe nach durch LC oder HPLC aufgetrennt. Die Marker werden von den aufgetrennten Fragmenten abgespalten und dann werden die Marker durch die jeweilige Nachweistechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen.

#### 15. Sequenzierung durch Hybridisierung

[0345] Die DNA-Sequenzanalyse wird herkömmlicherweise durch Hybridisieren eines Primers an Ziel-DNA und Durchführen von Kettenverlängerungen unter der Verwendung einer Polymerase durchgeführt. Spezifische Stops werden durch den Einbau eines Didesoxydnukleotids kontrolliert. Die Spezifität des Primers bei diesem Typ von Analyse kann durch Einschließen eines Hybotrops in dem Annealingpuffer und/oder in Kooperation eines abasischen Rests in dem Primer und Annealing bei einer unterscheidenden Temperatur erhöht werden.

[0346] Innerhalb einer Ausführungsform der Erfindung werden Verfahren zur Bestimmung der Identität eines Nukleinsäuremoleküls oder zum Nachweis eines ausgewählten Nukleinsäuremoleküls in, z. B., einer biologischen Probe zur Verfügung gestellt, wobei die allgemeinen Techniken der Sequenzierung durch Hybridisierung unter der Verwendung des Sanger-Verfahren verwendet werden. Kurz umfassen solche Verfahren im allgemeinen die Schritte von Hybridisieren eines markierten Primers an Ziel-DNA und Durchführung von Kettenverlängerungen unter der Verwendung einer Polymerase. Spezifische Stops werden durch den Einbau eines Didesoxynukleotids kontrolliert, das auch markiert sein kann. Die Zielnukleinsäuren werden gemäß ihrer Größe durch HPLC aufgetrennt. Die Marker werden von den aufgetrennten Fragmenten abgespalten und werden durch die jeweilige Nachweistechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen. Andere Sequenz-Analyseverfahren schließen die Hybridisierung des Ziels mit einer Auswahl von zufälligen, kurzen Oligonukleotiden ein. Die Sequenz wird durch überlappende Hybridisierungsanalyse konstruiert. Bei dieser Technik ist präzise Hybridisierung essentiell. Die Verwendung von hybotropen oder abasischen Resten und ein Annealing bei unterscheidenden Temperatur ist für dieses Verfahren vorteilhaft, um fehlgepaarte Hybridisierung zu verringern oder auszuschließen. Das Ziel ist es, automatisierte Hybridisierungsverfahren zu entwickeln, um große Arrays von Oligonukleotidsonden oder große Arrays von Nukleinsäureproben zu testen. Die Anwendungen von solchen Technologien schließen Genkartierung, Kloncharakterisierung, medizinische Genetik und Genentdeckung, DNA-Sequenzanalyse durch Hybridisierung und zuletzt Sequenzverifikation. Viele Parameter müssen kontrolliert werden, um Oligonukleotidsonden zu automatisieren oder einen Multiplex durchzuführen. Die Stabilität der jeweiligen Sonden muß ähnlich sein, das Ausmaß von Fehlpaarung mit der Zielnukleinsäure, die Temperatur, Ionenstärke, der A + T-Gehalt der Sonde (oder des Ziels) sowie andere Parameter wenn die Sonde kurz ist (d. h. 6–50 Nukleotide) sollten ähnlich sein. Gewöhnlicherweise werden die Bedingungen des Experiments und Sequenz der Sonde eingestellt, bis die Bildung der perfekten Basen-gepaarten Sonde thermodynamisch über jede Duplex, die eine Fehlpaarung enthält, bevorzugt wird. Die Anwendungen von Sonden in sehr großem Maßstab, wie z. B. bei Sequenzierung durch Hybridisierung (SBH) oder dem Testen von hochpolymorphen Loci, wie z. B. dem zystische Fibrose Trans-Membran Proteinlocus erfordern ein stringenteres Niveau der Kontrolle von Multiplex-Sonden. Innerhalb einer Ausführungsform der Erfindung werden Verfahren zum Nachweis der Identität eines Nukleinsäuremoleküls oder zum Nachweis eines ausgewählten Nukleinsäuremoleküls in, z. B. einer biologischen Probe zur Verfügung gestellt, welche die allgemeinen Techniken von Sequenzierung durch Hybridisierung anwenden. Kurz gesagt, umfassen solche Verfahren im allgemeinen ein Hybridisieren einer Reihe von markierten Primern an ein DNA-Ziel oder eine Reihe von Ziel-DNA-Fragmenten unter sorgfältig kontrollierten Bedingungen. Die Zielnukleinsäuren werden nach ihrer Größe durch HPLC aufgetrennt. Die Marker werden dann von den aufgetrennten Fragmenten abgespalten und durch die jeweilige Nachweistechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen.

#### 16. Oligonukleotid-Ligationsassay

[0347] Der Oligonukleotid-Ligationsassay ist eine Erweiterung von PCR-basiertem Screening, die einen ELI-

SA-basierten Assay verwendet (OLA, Nickerson et al., Proc. Natl. Acad. Sci. USA 87: 8923, 1990), um die PCR-Produkte nachzuweisen, die die Zielsequenz enthalten. Daher werden sowohl Gelelektrophorese und Koloniehybridisierung beseitigt. Kurz gesagt verwendet die OLA zwei angrenzende Oligonukleotide: eine „Reporter“-Sonde (markiert an 5'-Ende) und eine 5'-phosphorylierte/3'-biotinylierte „Verankerungs“-Sonde. Die zwei Oligonukleotide, die zu Sequenzen innerhalb der PCR-Primer komplementär sind, werden an Ziel-DNA annealed und, falls perfekte Komplementarität vorhanden ist, werden die zwei Sonden durch T4-DNA-Ligase ligiert. Ein Fangen der biotinylierten Verankerungs-sonde auf mobilisierten Streptavidin und die Analyse auf die kovalent gebundene Reportersonde testet auf die Anwesenheit oder Abwesenheit der Zielsequenzen unter den PCR-Produkten. Innerhalb einer Ausführungsform der Erfindung werden Verfahren zum Nachweis der Identität eines Nukleinsäuremoleküls oder zum Nachweis eines ausgewählten Nukleinsäuremoleküls, in, z. B., einer biologischen Probe zur Verfügung gestellt, die die Technik des Oligonukleotid-Ligationsassays verwenden. Kurz gesagt, umfassen solche Verfahren im allgemeinen die Schritte von Durchführen einer PCR auf der Ziel-DNA, gefolgt durch Hybridisierung mit der 5'-markierten „Reporter“-DNA-Sonde und einer 5'-phosphorylierten/nicht-biotinylierten Sonde. Die Probe wird mit T4-DNA-Ligase inkubiert. Die DNA-Stränge mit ligierten Proben können von der DNA mit nicht-ligierten Proben durch, z. B., bevorzugterweise LC oder HPLC abgetrennt werden. Die Marker werden dann von den abgetrennten Fragmenten abgespalten und dann werden die Marker durch die jeweilige Nachweistechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen. Kürzliche Fortschritte im OLA-Assay erlauben die gleichzeitige Analyse von multiplen Proben und multiplen Mutationen. (Baron et al., Nature Biotechnology 87: 1279, 1996). Kurz gesagt besteht das Verfahren aus der Amplifikation des Genfragments, das die Mutation von Interesse enthält, mit PCR. Das PCR-Produkt wird dann mit einer gewöhnlichen und zwei Allel-spezifischen Oligonukleotidsonden (wobei eine die Mutation enthält, während die andere dies nicht tut) hybridisiert, so daß die 3'-Enden der Allel-spezifischen Sonden unmittelbar angrenzend zu dem 5'-Ende der gewöhnlichen Sonde sind. Dieses erstellt einen kompetitiven Hybridisierungs-Ligationsprozeß zwischen den zwei allelischen Proben und der herkömmlichen Probe an jedem Locus. Die thermostabile DNA-Ligase unterscheidet dann zwischen Einzel-Base-Fehlpaarungen an der Verbindungsstelle, wodurch Allel-spezifische Ligationsprodukte erzeugt werden. Die herkömmliche Probe wird mit einem von vier Fluorophoren markiert und die Allel-spezifischen Sonden werden jede mit einem oder mehreren Pentaethylenoxid-Mobilitäts-modifizierenden Verlängerungen markiert, die einen Größenunterschied zwischen den verschiedenen Allel-spezifischen Sonden zur Verfügung stellen. Die Proben werden dann durch Gelelektrophorese basierend auf der Länge der modifizierenden Enden aufgetrennt und durch den fluoreszenten Marker an der herkömmlichen Sonde nachgewiesen. Durch die Verwendung in Größenunterschieden auf den Allel-spezifischen Sonden und vier für die herkömmliche Probe erhältlichen vier Fluorophore können viele Proben auf einer Spur des elektrophoretischen Gels analysiert werden. Innerhalb einer Ausführungsform der Erfindung werden Verfahren zum Nachweis der Identität eines Nukleinsäuremoleküls oder zum Nachweis eines ausgewählten Nukleinsäuremoleküls in, z. B. einer biologischen Probe zur Verfügung gestellt, wobei die Technik des Oligonukleotid-Ligationsassays für gleichzeitigen multiplen Probennachweis verwendet wird. Kurz gesagt, umfassen solche Verfahren allgemein die Schritte von Durchführung von PCR auf der markierten DNA, gefolgt von Hybridisierung mit der herkömmlichen Probe (nicht markiert) und den zwei Allel-spezifischen Sonden, die gemäß den Beschreibungen der Erfindung markiert sind. Die Probe wird mit DNA-Ligase inkubiert und die Fragmente, z. B. bevorzugterweise durch LC oder HPLC aufgetrennt. Die Marker werden von den aufgetrennten Fragmenten abgespalten und dann werden die Marker durch die jeweilige Nachweistechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen.

## 17. Differential Display

### a. Überblick

[0348] Säugetiere, wie z. B. der Mensch, weisen ungefähr 100.000 verschiedene Gene in ihrem Genom auf, von denen nur ein kleiner Teil, möglicherweise 15%, in jeder einzelnen Zelle exprimiert werden. Die Auswahl von exprimierten Genen bestimmt den biochemischen Charakter jeder bestimmten Zelle oder Gewebe. Das Verfahren von normalen zellulären Wachstum und Differenzierung sowie die pathologischen Veränderungen, die in Erkrankungen wie Krebs entstehen, werden alle durch Veränderungen in der Genexpression angetrieben. Differential-Display-Verfahren ermöglichen die Identifizierung von Genen, die spezifisch in einzelnen Zelltypen exprimiert werden.

[0349] Die Differential-Display-Technik amplifiziert die 3'-terminalen Teile von entsprechenden cDNAs durch Verwendung eines Primers, der so aufgebaut ist, um an die 5'-Grenze eines Poly(A)-Schwanzes zu binden und Primern mit beliebiger Sequenz, die stromaufwärts binden. Amplifizierte Populationen mit jedem Primerpaar werden durch ein Größentrennverfahren (PAGE HPLC, usw.) visualisiert, was einen direkten Vergleich der

mRNAs zwischen zwei biologischen Proben von Interesse ermöglicht. Das Differential-Display-Verfahren weist das Potential auf, um alle der exprimierten Gene (ungefähr 10.000 bis 15.000 mRNA Spezies) in einer Säugerzelle zu visualisieren und die Sequenzanalyse zu ermöglichen. Es ist möglich, zu vergleichen: (1) die Gesamtzahl von in den Eltern amplifizierten Peaks, (2) die Zahl von polymorphen Peaks zwischen Eltern und (3) die Segregationsverhältnisse von polymorphen Peaks in den Nachkommen von Kreuzungen von Tieren oder Pflanzen. Differential Display wird auch zur Identifizierung von Hoch- und Herunter-regulierten bekannten oder unbekannten Genen nach einer Vielzahl von Stimuli verwendet. Differential-Display-PCR-Fragmente können auch als Sonden für die cDNA-Klonierung verwendet werden (wobei ein unbekanntes Gen aus einer cDNA oder genomischen Bibliothek entdeckt wird).

[0350] Kurz gesagt sind die Schritte beim Differential-Display wie folgt: 1) RNA wird aus der biologischen Probe von Interesse isoliert. Gesamt-RNA, zytoplasmatische RNA oder mRNA kann verwendet werden. 2) Erster Strang cDNAs werden unter der Verwendung eines verankerten OligodT (OligodTdN, wobei N A, C oder G ist) erzeugt. 3) Amplifikation von cDNA unter der Verwendung von OligoTdN und kurzen Primern mit beliebiger Sequenz. Für eine vollständige Differential-Display-Analyse von zwei Zellpopulationen oder zwei Proben von Interesse werden 9 verschiedene Primer benötigt. Die Nachweisgrenze von Differential-Display für eine spezifische mRNA ist weniger als 0,001% der gesamten mRNA-Population.

[0351] Aufgrund der Einfachheit, Sensitivität und Reproduzierbarkeit des hier beschriebenen Verfahrens ist die CMST Differential-Display-Methode ein signifikanter Fortschritt über herkömmliche gelbasierte Systeme. Mit der CMST-basierten Differential-Display-Analyse von zwei Zelltypen können einschließlich 64 × 24 PCR-Läufe schnell vervollständigt werden, im Gegensatz zu einer arbeitsintensiven lang andauernden Zeit mit dem herkömmlichen Verfahren. Darüber hinaus wurde von Sequenzheterogenität von Banden, die aus Differential-Display-Gelen isoliert wurden gefunden, daß sie ein zu der hohen Fehlerrate dieser Technik beitragender Faktor sind. Dies wird mit der hier beschriebenen CMST-basierten Differential-Display-Methodologie vollständig vermieden.

#### b. CMST-basiertes Differential-Display Beispiel:

[0352] Das Ausgangsmaterial für Differential-Display ist RNA, die aus zwei verschiedenen Populationen von Zellen isoliert wurde. Im allgemeinen sind die Zellen von ähnlichem Ursprung und unterscheiden sich im Hinblick auf ihre Behandlung mit Wirkstoffen, ihrem „normalen“ gegenüber „transformierten“ Zustand oder ihrer Expression von verschiedenen eingefügten Genen.

[0353] Pflanzengewebe oder tierisches Material (2–3 g) wird geerntet und in sterile Petrischalen zerhackt oder zerschnitten. Das Material wird dann zu einem feinen Pulver in einem vorgekühlten Mörser- und Stöße unter flüssigem N<sub>2</sub> zerrieben. Ein Gramm von gefrorenem Pulver wird zu einem 12 ml Poly-Propylen-Röhrchen (1 g = ca. 5 ml Pulver), enthaltend 8 ml eines heißen (80°C) 1 : 1 Gemisches von RNA-Extraktionspuffer (100 mM Tris-HCl (pH 8,0), 100 mM LiCl, 10 mM EDTA, 1,0% LIDS) und Phenol (Basisvolumen.: 4 ml) transferiert. Die Probe wird bei hoher Geschwindigkeit für mindestens 30 Sekunden gemischt (gevortext). Ein Volumen von Chloroform wird hinzugefügt (4 ml) und nochmals gemischt und zentrifugiert für 20 Minuten in einer Zentrifuge bei 5000 bis 10.000 × g. Die wäßrige Phase wird in ein neues 12 ml Röhrchen transferiert und 1/3 Volumen von kaltem 8M LiCl wird hinzugefügt, um die RNA zu fällen (3 h bei 0°C). Die RNA wird bei 0°C für 20 Minuten zentrifugiert und in 1 ml H<sub>2</sub>O suspendiert. Restliche genomische DNA wird durch Behandlung der RNA-Probe mit DNase I beseitigt. Reverse Transkription wird durchgeführt auf jeder RNA-Probe unter der Verwendung von 500 ng DNA-freier RNA in 1x reverse Transkriptasepuffer, 10 mM DTT, 20 µM dNTPs, 0,2 µM 5'RS H-711 C (ein basenverankerter Primer mit 5'-Restriktionsstelle), 200 Einheiten MoMuLV reverser Transkriptase und 1,5 Einheiten RNA Guard pro 20 µl Reaktionsvolumen. Obwohl die Verwendung eines stromabwärts-Primers die Zahl von cDNA-Subfraktionen dreifach vermindert, reduziert sie nicht die Zahl von PCR-Reaktionen, die erforderlich sind, um die meisten der cDNA-Spezies, die in dem Pool vorhanden sind, darzustellen. Im Gegenteil verringert sie die theoretische Chance, diejenigen cDNA-Spezies zu identifizieren, die vorhanden sind. Die besten Ergebnisse werden unter der Verwendung einer Kombination von neun verschiedenen Primern des Typs DMO-VV, wobei V A, G, C sein kann, jedoch nicht T, erhalten. Mit einem T in der terminalen 3'-Position führt die unvollständige Hybridisierung des Primers zu einem Verschmieren von Banden auf dem Gel. Die optimale Konzentration von RNA ist 200–300 ng pro cDNA-Synthese.

[0354] Das CMST-basierte Differential-Display wird im wesentlichen wie vorher beschrieben (Liang and Pardee, Science 257: 967–971, 1992) durchgeführt, mit der Ausnahme für den Aufbau von für die reverse Transkription und die Amplifikationsschritte verwendeten Primer und die Auswahl von radiomarkiertem Nukleotid. Eine vollständige Differential-Display-Analyse der cDNAs aus zwei biologischen Proben von Interesse unter der Verwendung von neun Stromabwärts-Primern und 24 Stromaufwärts-Primern würde 9 × 24 × 2 CMST-basierte Differential-Display-Reaktionen erzeugen. Die Amplifikationsprodukte können durch HPLC aufgetrennt werden und, falls gewünscht, reamplifiziert werden.

[0355] Im Anschluß an die Inkubation der RNA bei 65°C für 5 Minuten werden die Proben auf Eis gekühlt, zu

den Reverse-Transkriptionsmix hinzugefügt und für 60 Minuten bei 37°C inkubiert, gefolgt von 95°C für 5 Minuten. Duplikat cDNA-Proben werden dann unter der Verwendung desselben 5'-Primers in Kombination mit einer Reihe von 13-Meren von beliebiger, jedoch definierter Sequenz: H-AP: AAGCTTCGACTGT, H-AP: AAGCTTTGGTCAG, H-AP4: AAGCTTCTCAACG, H-AP5: AAGCTTAGTAGGC amplifiziert.

[0356] Die Modifikation wird in Reaktionsgemischen durchgeführt, die 0,1 Volumen von reverser Transkriptionsreaktion enthalten, 1 × PCR-Puffer (10 × PCR-Puffer = 100 mM Tris-HCl, 15 mM MgCl<sub>2</sub>, 10 mM KCl, pH 8,3) 2 µM dNTPs, 0,2 µM RS H-T11C verankertem Primer, 0,2 µM geeignetem beliebigem Primer, 1,5 Einheiten Expand™ High Fidelity DNA-Polymerase und Wasser bis zu einem finalen Volumen von 20 µl. Die Amplifikation der cDNA wird unter den folgenden Bedingungen durchgeführt: 94°C (1 Minute), gefolgt von 40 Zyklen von 94°C (30 Sekunden), 40°C (2 Minuten), 72°C (30 Sekunden) und beendet mit 72°C (5 Minuten).

[0357] Die Amplifikation für jedes Gen wird mit genspezifischen Primern durchgeführt, die eine bekannte Intron/Exon-Grenze überspannen (siehe unten). Alle Amplifikationen werden in 20 µl Volumen durchgeführt, die 10 mM Tris HCl pH 8,3, 1 mM NH<sub>4</sub>Cl, 1,5 M MgCl<sub>2</sub>, 100 mM KCl, 0,125 mM NTPs, 10 ng/ml der jeweiligen Oligonukleotidprimer und 0,75 Einheiten von Taq-DNA-Polymerase (Gibco-BRL) enthalten. Die Zyklusparameter waren 94°C Vorheisschritt für 5 Minuten, gefolgt von 94°C Denaturierungsschritt für 1 Minute, 55°C Annealingschritt für 2 Minuten und einem 72°C Verlängerungsschritt für 30 Sekunden bis 1 Minute und einer finalen Verlängerung von 72°C für 10 Minuten. Die Zahl der Amplifikationszyklen lag im allgemeinen von 30–45.

[0358] Die Amplifikationsprodukte werden Gel-gereinigt (Zhen and Swank, BioTechniques 14: 894-898, 1993) auf 1% Agarosegelen, die in 0,04 M Tris-Acetat, 0,001 M EDTA (1 × TEA)-Puffer laufen gelassen wurden und mit Ethidiumbromid gefärbt wurden. Eine Mulde wird gerade vor der Bande von Interesse eingeschnitten und mit 50–200 µl von 10% PEG in 1 × TAE-Puffer gefüllt. Die Elektrophorese wird fortgeführt, bis die Bande vollständig in die Mulde eingetreten ist. Die Inhalte werden dann entfernt und mit Phenol, Chloroform extrahiert und in 0,1 Volumen von 7,5 M Ammoniumacetat und 2,5 Volumen 100% EtOH gefällt. Die Proben werden mit 75% EtOH gewaschen und kurz bei Raumtemperatur getrocknet. Die Quantifizierung der Ausbeute wird durch Elektrophorese eines kleinen Aliquots auf 1% Agarosegel in 1 × TBE-Puffer mit Ethidiumbromidfärbung und Vergleich mit einem bekannten Standard durchgeführt.

[0359] Die Produkte aus den Amplifikationsreaktionen werden durch HPLC analysiert. Die HPLC wird unter der Verwendung von automatisierter HPLC-Instrumentierung (Rainin, Emeryville, CA, oder Hewlett Packard, Palo Alto, CA) durchgeführt. Ungereinigte DNA-Fingerprinting-Produkte, die für 3 Minuten bei 95°C vor der Injektion in eine HPLC denaturiert werden, werden mit linearen Acetonitril (ACN, J. T. Baker, NJ) Gradienten von 1,8%/Minute bei einer Flußrate von 0,9 ml/Minute eluiert. Die Start- und Endpunkte werden gemäß der Größe der amplifizierten Produkte eingestellt. Die für die erfolgreiche Auflösung der Moleküle, die während der DNA-Fingerprinting-Technik erzeugt wurden erforderliche Temperatur ist 50°C.

[0360] Der Efflux aus der HPLC wird dann in einen Massenspektrometer (Hewlett Packard, Palo Alto, CA) zum Nachweis der Marker umgeleitet.

[0361] Der Vergleich der Chromatogramme (Massenspektrometrie-basiert) zeigt, daß Banden bei 220 bp und 468 bp in den stimulierten Jurkat-Zellen beobachtet werden und nicht in den nicht-stimulierten Jurkat-Zellen beobachtet werden.

## C. TRENNUNG VON NUKLEINSÄUREFRAGMENTEN

[0362] Eine Probe, die eine Analyse erfordert, ist oft eine Mischung aus vielen Komponenten in einer komplexen Matrix. Für Proben, die unbekannte Verbindungen enthalten, müssen die Komponenten voneinander getrennt werden, so daß jede einzelne Komponente mit anderen analytischen Verfahren identifiziert werden kann. Die Trenneigenschaften der Komponenten in einer Mischung sind unter konstanten Bedingungen konstant, und wenn sie daher erst einmal bestimmt sind, können sie verwendet werden, um jede der Komponenten zu identifizieren und zu quantifizieren. Solche Verfahren sind typisch bei chromatographischen oder elektrophoretischen analytischen Trennungen.

### 1. Hochleistungsflüssigkeitschromatographie (HPLC)

[0363] Hochleistungsflüssigkeitschromatographie (HPLC) ist eine chromatographische Trenntechnik, um Verbindungen zu trennen, die in Lösung gelöst sind. HPLC-Instrumente bestehen aus einem Reservoir aus mobiler Phase, einer Pumpe, einem Injektor, einer Trennsäule und einem Detektor. Verbindungen werden durch Injizieren eines aliquoten Teils der Probenmischung auf die Säule getrennt. Die unterschiedlichen Verbindungen gehen aufgrund der Unterschiede in ihrem Verteilungsverhalten zwischen der mobilen flüssigen Phase und der stationären Phase mit unterschiedlichen Geschwindigkeiten durch die Säule hindurch.

[0364] Vor kurzem ist gezeigt worden, daß IP-RO-HPLC auf nicht-porösen PS/DVB-Teilchen mit chemisch gebundenen Alkylketten schnelle Alternativen zu Kapillarelektrophorese bei der Analyse von sowohl einzel- als auch doppelsträngigen Nukleinsäuren ist, die ähnliche Auflösungsgrade liefert (Huber et al., Anal. Biochem.

212: 351, 1993; Huber et al., 1993, Nuc. Acids Res. 21: 1061; Huber et al, Biotechniques 16: 898, 1993). Im Gegensatz zu Ionenaustauschchromatographie, die nicht immer doppelsträngige DNA als eine Funktion der Stranglänge zurückhält (da AT-Basenpaare mit der positiv geladenen stationären Phase stärker als GC-Basenpaare wechselwirken), ermöglicht IP-RP-HPLC eine streng größenabhängige Trennung.

[0365] Ein Verfahren unter Verwendung von 100 mM Triethylammoniumacetat als Ionenpaarungsreagens ist entwickelt worden, Phosphodiester-Oligonukleotide konnten mit Erfolg auf alkylierten nicht-porösen 2,3 µM Poly(styrol-divinylbenzol)-Teilchen mittels Hochleistungsflüssigkeitschromatographie getrennt werden (Oefner et al., Anal. Biochem. 223: 39, 1994). Die beschriebene Technik ermöglichte die Trennung von PCR-Produkten, die sich nur 4 bis 8 Basenpaare in der Länge unterscheiden, in einem Größenbereich von 50 bis 200 Nukleotiden.

## 2. Elektrophorese

[0366] Elektrophorese ist eine Trenntechnik, die auf der Mobilität von Ionen (oder DNA, wie dies der hierin beschriebene Fall ist) in einem elektrischen Feld beruht. Negativ geladene DNA wandert zu einer positiven Elektrode und positiv geladene Ionen wandern zu einer negativen Elektrode. Aus Sicherheitsgründen ist üblicherweise eine Elektrode geerdet und die andere positiv oder negativ vorgespannt. Geladene Spezies haben unterschiedliche Wanderungsgeschwindigkeiten in Abhängigkeit von ihrer Gesamtladung, Größe und Gestalt und können daher getrennt werden. Ein Elektrodenapparat besteht aus einer Hochspannungs-Stromquelle, Elektroden, Puffer und einem Träger für den Puffer, wie etwa Polyacrylamidgel, oder eine Kapillarröhre. Offene Kapillarröhren werden für viele Arten von Proben verwendet und die anderen Gelträger werden üblicherweise für biologische Proben verwendet, wie etwa Proteinmischungen oder DNA-Fragmente.

## 3. Kapillarelektrophorese (CE)

[0367] Kapillarelektrophorese (CE) in ihren verschiedenen Manifestationen (freie Lösung, Isotachophorese, isoelektrische Fokussierung, Polyacrylamidgel, mizelläre elektrokinetische "Chromatographie") entwickelt sich als ein Verfahren für schnelle Hochauflösungstrennungen sehr kleiner Probenvolumina komplexer Mischungen. In Kombination mit der inhärenten Empfindlichkeit und Selektivität von MS ist CE-MS eine potentielle leistungsfähige Technik für die Bioanalyse. In der hierin offenbarten neuartigen Anwendung wird das Zusammenspiel dieser zwei Verfahren zu überlegenen DNA-Sequenzierungsverfahren führen, die die Sequenzierungsverfahren mit der augenblicklichen Geschwindigkeit um mehrere Größenordnungen übertreffen.

[0368] Die Beziehung zwischen CE und Durchflüssen bei Elektrosprayionisierung (ESI) und die Tatsache, daß beide durch Ionenspezies in Lösung erleichtert werden (und primär dafür verwendet werden), liefert die Grundlage für eine extrem attraktive Kombination. Die Kombination von sowohl Kapillarzonenelektrophorese (CZE) und Kapillar isotachophorese mit Quadrupol-Massenspektrometern auf der Basis von ESI ist beschrieben worden (Olivares et al., Anal. Chem. 59: 1230, 1987; Smith et al., Anal. Chem. 60: 436, 1988; Loo et al., Anal. Chem. 179: 404, 1989; Edmonds et al., J. Chroma. 474: 21, 1989; Loo et al., J. Microcolumn Sep. 1: 223, 1989; Lee et al., J. Chromatog. 458: 313, 1988; Smith et al., J. Chromatog. 480: 211, 1989; Grese et al., J. Am. Chem. Soc. 111: 2835, 1989). Kleine Peptide sind einer CZE-Analyse mit guter Empfindlichkeit (im Femtomol-Bereich) leicht zugänglich.

[0369] Das leistungsfähigste Trennverfahren für DNA-Fragmente ist Polyacrylamidgelelektrophorese (PAGE), im allgemeinen in einem Slabgel-Format. Die wesentliche Beschränkung der gegenwärtigen Technologie ist jedoch die relativ lange Zeit, die erforderlich ist, um die Gelelektrophorese von DNA-Fragmenten, die in den Sequenzierungsreaktionen hergestellt werden, durchzuführen. Eine Erhöhung der Größenordnung (10-fach) kann mit der Verwendung von Kapillarelektrophorese erreicht werden, die ultradünne Gele einsetzt. In freier Lösung wandert, in einer ersten Annäherung, die gesamte DNA mit derselben Mobilität, da die Zugabe einer Base zur Kompensation von Masse und Ladung führt. In Polyacrylamidgelen sieben und wandern DNA-Fragmente als eine Funktion der Länge, und dieser Ansatz ist nunmehr auf CE angewendet worden. Bemerkenswerte Bodenzahlen pro Meter sind jetzt mit quervernetztem Polyacrylamid erreicht worden ( $10^{+7}$  Böden pro Meter, Cohen et al., Proc. Natl. Acad. Sci., USA 85: 9660, 1988). Solche CE-Säulen, wie beschrieben, können zur DNA-Sequenzierung eingesetzt werden. Das CE-Verfahren ist im Prinzip 25-mal schneller als Slabgel-Elektrophorese in einem Standard-Sequencer. Zum Beispiel können etwa 300 Basen pro Stunde gelesen werden. Die Trennungsgeschwindigkeit ist bei Slabgel-Elektrophorese durch die Größenordnung des elektrischen Feldes beschränkt, das ohne übermäßige Wärmeproduktion an das Gel angelegt werden kann. Daher wird die größere Geschwindigkeit von CE durch die Verwendung höherer Feldstärken erreicht (300 V/cm bei CE gegenüber 10 V/cm bei Slabgel-Elektrophorese). Das Kapillarformat verringert den Stromfluß und damit die Leistung und die resultierende Wärmezeugung.

[0370] Smith und andere (Smith et al., Nuc. Acids Res. 18: 4417, 1990) haben den Einsatz mehrerer Kapillaren in Parallelschaltung vorgeschlagen, um den Durchsatz zu erhöhen. In ähnlicher Weise haben Mathies und

Huang (Mathies und Huang, *Nature* 359: 167, 1992) Kapillarelektrophorese eingeführt, bei der Trennungen auf einer Parallelanordnung von Kapillaren durchgeführt werden, und zeigten Sequenzierung mit hohem Durchsatz (Huang et al., *Anal. Chem.* 64: 967, 1992, Huang et al., *Anal. Chem.* 64: 2149, 1992). Der wichtigste Nachteil von Kapillarelektrophorese ist die begrenzte Probenmenge, die auf die Kapillare geladen werden kann. Durch Konzentration einer großen Probenmenge am Beginn der Kapillare, vor der Trennung, wird die Beladbarkeit erhöht, und Nachweisgrenzen können um mehrere Größenordnungen gesenkt werden. Das populärste Verfahren zur Vorkonzentration bei CE ist Probenstapelung. Probenstapelung ist vor kurzem in einer Übersicht dargestellt worden (Chien und Burgi, *Anal. Chem.* 64: 489A, 1992). Probenstapelung hängt ab von dem Matrixunterschied (pH, Ionenstärke) zwischen dem Probenpuffer und dem Kapillarpuffer, so daß das elektrische Feld über die Probenzone stärker ist als in der Kapillarregion. Bei Probenstapelung wird ein großes Probenvolumen in einem Puffer mit geringer Konzentration zur Vorkonzentrierung am Kopf der Kapillarsäule eingeführt. Die Kapillare wird mit einem Puffer derselben Zusammensetzung gefüllt, aber mit höherer Konzentration. Wenn die Proben-Ionen den Kapillarpuffer und das niedrigere elektrische Feld erreichen, stapeln sie sich in eine konzentrierte Zone hinein. Probenstapelung hat die Nachweisbarkeiten um 1–3 Größenordnungen erhöht.

[0371] Ein weiteres Verfahren zur Vorkonzentration ist die Anwendung von Isotachophorese (ITP) vor der Freizonen-CE-Trennung von Analyten. ITP ist eine Elektrophoresetechnik, die es ermöglicht, daß Mikroliter-Probenvolumina auf die Kapillare geladen werden können, im Gegensatz zu den Injektionsvolumina von wenigen nl, die typischerweise mit CE assoziiert sind. Die Technik beruht auf der Einführung der Probe zwischen zwei Puffern (Leading- und Trailing-Elektrolyten) mit höherer bzw. niedrigerer Mobilität als der Analyt. Die Technik ist inhärent eine Konzentrationstechnik, bei der die Analyten sich in reine Zonen hineinkonzentrieren, die mit derselben Geschwindigkeit wandern. Die Technik ist gegenwärtig weniger populär als die Stapelverfahren, die oben beschrieben sind, wegen des Bedarfs mehrerer Wahlmöglichkeiten von Leading- und Trailing-Elektrolyten und der Fähigkeit, nur kationische oder anionische Spezies während eines Trennverfahrens zu trennen.

[0372] Das Herz des DNA-Sequenzierungsverfahrens ist die bemerkenswert selektive elektrophoretische Trennung von DNA- oder Oligonukleotid-Fragmenten. Sie ist bemerkenswert, weil jedes Fragment aufgelöst wird und sich nur um Nukleotide unterscheidet. Trennungen von bis zu 1.000 Fragmenten (1.000 bp) sind erzielt worden. Ein weiterer Vorteil der Sequenzierung mit spaltbaren Markierungen ist wie folgt. Es gibt kein Erfordernis, ein Slabgel-Format zu verwenden, wenn DNA-Fragmente mit Polyacrylamidgelelektrophorese getrennt werden, wenn spaltbare Markierungen eingesetzt werden. Da zahlreiche Proben kombiniert werden (4 bis 2.000), besteht keine Notwendigkeit, Proben parallel laufen zu lassen, wie dies der Fall bei gegenwärtigen Farbstoff-Primer- oder Farbstoff-Terminator-Verfahren ist (d. h. ABI373-Sequencer). Da es keinen Grund gibt, Parallelspuren laufen zu lassen, gibt es keinen Grund, ein Slabgel zu verwenden. Daher kann man ein Röhrenformat für das Elektrophoresetrennverfahren einsetzen. Grossman (Grossman et al., *Genet. Anal. Tech. Appl.* 9: 9, 1992) hat gezeigt, daß ein beträchtlicher Vorteil gewonnen wird, wenn ein Röhrenformat anstelle eines Slabgel-Formats verwendet wird. Dies beruht auf der größeren Fähigkeit, Joule-Wärme in einem Röhrenformat abzuleiten, verglichen mit einem Slabgel, was zu schnelleren Laufzeiten führt (um 50%) und viel höherer Auflösung von DNA-Fragmenten mit hoher relativer Molekülmasse (mehr als 1.000 nt). Lange Ableitungen sind bei genomischer Sequenzierung entscheidend. Daher hat die Verwendung spaltbarer Markierungen beim Sequenzieren den zusätzlichen Vorteil, daß sie dem Benutzer erlaubt, das effizienteste und empfindlichste DNA-Trennverfahren einzusetzen, das auch die höchste Auflösung besitzt.

#### 4. Mikrosysteme

[0373] Kapillarelektrophorese (CE) ist ein leistungsfähiges Verfahren zur DNA-Sequenzierung, forensischen Analysen, PCR-Produktanalyse und Restriktionsfragmentgrößenbestimmung. CE ist weit schneller als traditionelle Slab-PAGE, da bei Kapillargelen ein weit höheres Spannungsfeld angelegt werden kann. CE hat jedoch den Nachteil, daß es die Verarbeitung von nur einer Probe pro Gel erlaubt. Das Verfahren kombiniert die schnelleren Trennzeiten von CE mit der Fähigkeit, mehrere Proben parallel zu analysieren. Das zugrundeliegende Konzept hinter der Verwendung von Mikrosystemen ist die Fähigkeit, die Informationsdichte in der Elektrophorese durch Miniaturisierung der Spwabmessung auf etwa 100 Mikrometer zu erhöhen. Die Elektronikindustrie verwendet routinemäßig Mikrotechnologie, um Schaltungen mit Merkmalen von weniger als 1 Mikron in der Größe herzustellen. Die gegenwärtige Dichte von Kapillaranordnungen ist beschränkt durch den Außendurchmesser der Kapillarröhre. Mikrotechnologisch hergestellte Kanäle erzeugen eine höhere Dichte von Anordnungen. Mikrotechnologie ermöglicht auch physikalische Baugruppen, die mit Glasfasern nicht möglich sind, und verbindet die Kanäle direkt mit anderen Vorrichtungen auf einem Chip. Mehrere Vorrichtungen sind auf Mikrochips für Trenntechnologien konstruiert worden. Ein Gaschromatograph (Teny et al., *IEEE Trans. Electron Device*, ED-26: 1880, 1979) und ein Flüssigkeitschromatograph (Manz et al., *Sens. Actuators B* 1: 249, 1990) sind auf Siliciumchips hergestellt worden, aber diese Vorrichtungen sind nicht in breitem Umfang eingesetzt worden. Mehrere Gruppen haben über die Trennung von Fluoreszenzfarbstoffen und Aminosäuren



auf mikrotechnologisch hergestellten Vorrichtungen berichtet (Manz et al., J. Chromatography 593: 253, 1992, Effenhauser et al., Anal. Chem. 65: 2637, 1993). Vor kurzem haben Woolley und Mathies (Woolley und Mathies, Proc. Natl. Acad. Sci. 91: 11348, 1994) gezeigt, daß Photolithographie und chemisches Ätzen verwendet werden können, um eine große Anzahl von Trennkanälen auf Glassubstraten herzustellen. Die Kanäle werden mit Hydroxyethylcellulose(HEC)trennmatrixen gefüllt. Es wurde gezeigt, daß DNA-Restriktionsfragmente in so wenig wie zwei Minuten getrennt werden konnten.

#### D. ABSPALTUNG VON MARKIERUNGEN

[0374] Wie oben beschrieben, werden unterschiedliche Linkerkonstruktionen Spaltbarkeit ("Labilität") unter unterschiedlichen spezifischen physikalischen oder chemischen Bedingungen verleihen. Beispiele für Bedingungen, die dazu dienen, verschiedene Linkerkonstruktionen zu spalten, schließen saure, basische, Oxidations-, Reduktions-, Fluorid-, Thiolaustausch-, Photolyse- und enzymatische Bedingungen ein.

[0375] Beispiele für spaltbare Linker, die die allgemeinen Kriterien für Linker, die oben aufgelistet sind, erfüllen, werden den Fachleuten gut bekannt sein und schließen diejenigen ein, die im von Pierce (Rockford, IL) erhältlichen Katalog zu finden sind. Beispiele schließen ein:

- Ethylenglycobis(succinimidylsuccinat) (EGS), ein aminreaktives Vernetzungsreagens, das mit Hydroxylamin spaltbar ist (1 M bei 37°C für 3–6 Stunden);
- Disuccinimidyltartrat (DST) und Sulfo-DST, die aminreaktive Vernetzungsreagentien sind, spaltbar mit 0,015 M Natriumperiodat;
- Bis[2-(succinimidylsuccinoyloxy)ethyl]sulfon (BSOCOES) und Sulfo-BSOCOES, die aminreaktive Vernetzungsmittel sind, spaltbar durch Base (pH 11,6);
- 1,4-Di[3'-(2'-pyridyldithio(propionamido))butan] (DPDPB), ein Pyridyldithiol-Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist;
- N-[4-(p-Azidosalicylamido)-butyl]-3'-(2'-pyridyldithio)propionamid (APDP), ein Pyridyldithiol-Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist;
- Bis-[beta-4-(azidosalicylamido)ethyl]-disulfid, ein photoreaktiver Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist;
- N-Succinimidyl-(4-azidophenyl)-1,3'-dithiopropionat (SADP), ein photoreaktiver Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist;
- Sulfosuccinimidyl-2-(7-azido-4-methylcumarin-3-acetamid)ethyl-1,3'-dithiopropionat (SAED), ein photoreaktiver Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist;
- Sulfosuccinimidyl-2-(m-azido-o-nitrobenzamido)-ethyl-1,3'-dithiopropionat (SAND), ein photoreaktiver Vernetzer, der durch Thiolaustausch oder Reduktion spaltbar ist.

[0376] Andere Beispiele für spaltbare Linker und die Spaltungsbedingungen, die verwendet werden können, um Markierungen freizusetzen, sind wie folgt. Eine Silyl-Verknüpfungsgruppe kann durch Fluorid oder unter sauren Bedingungen gespalten werden. Eine 3-, 4-, 5- oder 6-substituierte 2-Nitrobenzyloxy- oder 2-, 3-, 5- oder 6-substituierte 4-Nitrobenzyloxy-Verknüpfungsgruppe kann durch eine Photonenquelle (Photolyse) gespalten werden. Eine 3-, 4-, 5- oder 6-substituierte 2-Alkoxyphenoxy- oder 2-, 3-, 5- oder 6-substituierte 4-Alkoxyphenoxy-Verknüpfungsgruppe kann durch  $\text{Ce}(\text{NH}_4)_2(\text{NO}_3)_6$  (Oxidation) gespalten werden. Ein  $\text{NCO}_2$  (Urethan)-Linker kann durch Hydroxid (Base), Säure oder  $\text{LiAlH}_4$  (Reduktion) gespalten werden. Eine 3-Pentenyl-, 2-Butenyl- oder 1-Butenyl-Verknüpfungsgruppe kann durch  $\text{O}_3$ ,  $\text{OsO}_4/\text{IO}_4^-$  oder  $\text{KMnO}_4$  (Oxidation) gespalten werden. Eine 2-[3-, 4- oder 5-substituierte Furyl]oxy-Verknüpfungsgruppe kann durch  $\text{O}_2$ ,  $\text{Br}_2$ ,  $\text{MeOH}$  oder Säure gespalten werden.

[0377] Bedingungen für die Spaltung weiterer labiler Verknüpfungsgruppen schließen ein: tert.-Alkyloxy-Verknüpfungsgruppen können durch Säure gespalten werden; Methyl(dialkyl)methoxy- oder 4-substituierte 2-Alkyl-1,3-dioxan-2-yl-Verknüpfungsgruppen können durch  $\text{H}_3\text{O}^+$  gespalten werden; 2-Silylethoxy-Verknüpfungsgruppen können durch Fluorid oder durch Säure gespalten werden; 2-(X)-Ethoxy (wobei X = Keto, Esteramid, Cyano,  $\text{NO}_2$ , Sulfid, Sulfoxid, Sulfon ist)-Verknüpfungsgruppen können unter alkalischen Bedingungen gespalten werden; 2-, 3-, 4-, 5- oder 6-substituierte Benzyloxy-Verknüpfungsgruppen können durch Säure oder unter reduktiven Bedingungen gespalten werden; 2-Butenyl-Verknüpfungsgruppen können durch  $(\text{Ph}_3\text{P})_3\text{RhCl}(\text{H})$  gespalten werden; 3-, 4-, 5- oder 6-substituierte 2-Bromphenoxy-Verknüpfungsgruppen können durch Li, Mg oder BuLi gespalten werden; Methylthiomethoxy-Verknüpfungsgruppen können durch  $\text{Hg}^{2+}$  gespalten werden; 2-(X)-Ethoxy (wobei X = ein Halogen)-Verknüpfungsgruppen können durch Zn oder Mg gespalten werden; 2-Hydroxyethoxy-Verknüpfungsgruppen können durch Oxidation (z. B. mit  $\text{Pb}(\text{OAc})_4$ ) gespalten werden.

[0378] Bevorzugte Linker sind diejenigen, die durch Säure oder Photolyse gespalten werden. Mehrere der säurelabilen Linker, die für Festphasen-Peptidsynthese entwickelt worden sind, sind zur Verknüpfung von Markierungen mit MOIs nützlich. Einige dieser Linker sind bei einem kürzlich erschienenen Überblick von

Lloyd-Williams et al. beschrieben (Tetrahedron 49: 11065–11133, 1993). Ein nützlicher Linker-Typ beruht auf p-Alkoxybenzylalkoholen, von denen zwei, 4-Hydroxymethylphenoxyessigsäure und 4-(4-Hydroxymethyl-3-methoxyphenoxy)buttersäure, von Advanced ChemTech (Louisville, KY) kommerziell erhältlich sind. Beide Linker können über eine Esterbindung mit dem Benzylalkohol an eine Markierung gebunden werden und über eine Amidbindung mit der Carbonsäure an ein aminhaltiges MOI. Markierungen, die durch diese Moleküle gebunden sind, werden vom MOI mit variierenden Konzentrationen in Trifluoressigsäure freigesetzt. Die Spaltung dieser Linker führt zur Freisetzung einer Carbonsäure auf der Markierung. Säurespaltung von Markierungen, die durch verwandte Linker gebunden sind, wie etwa 2-Dimethoxy-4'-(carboxymethoxy)-benzhydrylamin (erhältlich von Advanced ChemTech in Fmoc-geschützter Form), führt zur Freisetzung eines Carbonsäureamids auf der freigesetzten Markierung.

[0379] Die photolabilen Linker, die für diese Anwendung nützlich sind, sind ebenfalls größtenteils für die Festphasen-Peptidsynthese entwickelt worden (siehe Übersicht von Lloyd-Williams). Diese Linker beruhen üblicherweise auf 2-Nitrobenzylestern oder 2-Nitrobenzylamiden. Zwei Beispiele für photolabile Linker, über die vor kurzem in der Literatur berichtet worden ist, sind 4-(4-(1-Fmoc-Amino)ethyl)-2-methoxy-5-nitrophenoxy)butansäure (Holmes und Jones, J. Org. Chem. 60: 2318–2319, 1995) und 3-(Fmoc-Amino)-3-(2-nitrophenyl)propionsäure (Brown et al., Molecular Diversity 1: 4–12, 1995). Beide Linker können über die Carbonsäure an ein Amin auf dem MOI gebunden werden. Die Bindung der Markierung an den Linker erfolgt durch Bildung eines Amids zwischen einer Carbonsäure auf der Markierung und dem Amin auf dem Linker. Die Spaltung photolabiler Linker erfolgt üblicherweise mit UV-Licht mit 350 nm Wellenlänge bei Intensitäten und Zeiträumen, die den Fachleuten bekannt sind. Die Spaltung der Linker führt zu einer Freisetzung eines primären Amids auf der Markierung. Beispiele für photospaltbare Linker schließen Nitrophenylglycinester, exo- und endo-2-benzonorborenylchloride und Methansulfonate und 3-Amino-3(2-nitrophenyl)propionsäure ein. Beispiele für enzymatische Spaltung schließen Esterasen, die Esterbindungen spalten werden, Nukleasen, die Phosphodiester spalten werden, Proteasen, die Peptidbindungen spalten werden, etc. ein.

[0380] Geeignete Vorrichtungen, die dazu verwendet werden können, die Photospaltung von markierten Molekülen durchzuführen, schließen die Vorrichtung ein, die unter der Abkürzung „PHRED“ bekannt ist, was für Photochemischer Reaktor für den verstärkten Nachweis steht und von Aura Industries, Staten Island, NY erhältlich ist (erhältlich mit sowohl einer 254 nm und 366 nm Lampe), PhotoBlaster System –1 mit dem LuxTube Aufbau, von Agrenetics, 81 Salem Street, Wilmington, MA, USA 01187 (erhältlich mit einer 366 nm Lampe, jedoch enthält sie Röhre einen Photokatalysator, der durch 366 nm Licht aktiviert wird, was zu einer Emission von einem Bereich von Wellenlängen führt, der 254 nm einschließt).

## E. NACHWEIS VON MARKIERUNGEN

[0381] Nachweisverfahren beruhen typischerweise auf der Absorption und Emission in bestimmten Spektraltypen. Wenn Atome oder Moleküle Licht absorbieren, regt die einstrahlende Energie eine quantisierte Struktur in ein höheres Energieniveau an. Der Anregungstyp hängt von der Wellenlänge des Lichtes ab. Elektronen werden durch ultraviolettes oder sichtbares Licht auf höhere Orbitale angehoben, Molekülvibrationen werden durch Infrarotlicht angeregt und Rotationen werden durch Mikrowellen angeregt. Ein Absorptionsspektrum ist die Lichtabsorption als eine Funktion der Wellenlänge. Das Spektrum eines Atoms oder Moleküls hängt von seiner Energieniveaustuktur ab. Absorptionsspektren sind nützlich für die Identifizierung von Verbindungen. Spezifische absorptionspektroskopische Verfahren schließen Atomabsorptionsspektroskopie (AA), Infrarotspektroskopie (IR) und UV-vis-Spektroskopie (UV-vis) ein.

[0382] Atome oder Moleküle, die auf hohe Energieniveaus angeregt werden, können durch Abgabe von Strahlung auf niedrigere Niveaus absteigen. Diese Lichtemission wird Fluoreszenz genannt, wenn der Übergang zwischen Zuständen mit demselben Spin erfolgt, und Phosphoreszenz, wenn der Übergang zwischen Zuständen mit unterschiedlichem Spin auftritt. Die Emissionsintensität des Analyten ist linear proportional zur Konzentration (bei niedrigen Konzentrationen) und ist nützlich zur Quantifizierung der emittierenden Spezies. Spezifische emissionspektroskopische Verfahren schließen Atomemissionsspektroskopie (AES), Atomfluoreszenzspektroskopie (AFS), molekulare laserinduzierte Fluoreszenz (LIF) und Röntgenfluoreszenz (XRF) ein.

[0383] Wenn elektromagnetische Strahlung durch die Materie hindurchgeht, setzt der größte Teil der Strahlung seinen Weg in seiner ursprünglichen Richtung fort, ein kleinerer Teil wird aber in andere Richtungen gestreut. Licht, das mit derselben Wellenlänge wie das einstrahlende Licht gestreut wird, wird Rayleigh-Streuung genannt. Licht, das in durchsichtigen Feststoffen aufgrund von Vibrationen (Phononen) gestreut wird, wird Brillouin-Streuung genannt. Brillouin-Streuung ist typischerweise um 0,1 bis 1 Wellenzahl vom einstrahlenden Licht verschoben. Licht, das aufgrund von Vibrationen in Molekülen oder optischen Phononen in undurchsichtigen Feststoffen gestreut wird, wird Raman-Streuung genannt. Raman-gestreutes Licht wird um so viel wie 4.000 Wellenzahlen vom einstrahlenden Licht verschoben. Spezifische streuungsspektroskopische Verfahren schließen Raman-Spektroskopie ein.

[0384] IR-Spektroskopie ist die Messung der Wellenlänge und Intensität der Absorption von Licht im mittleren Infrarotbereich durch eine Probe. Licht im mittleren Infrarotbereich ( $2,5\text{--}50\text{ }\mu\text{m}$ ,  $4.000\text{--}200\text{ cm}^{-1}$ ) ist energetisch ausreichend, um Molekylvibrationen auf höhere Energieniveaus anzuregen. Die Wellenlänge der IR-Absorptionsbanden ist charakteristisch für spezifische Typen von chemischen Bindungen und IR-Spektroskopie ist im allgemeinen am nützlichsten zur Identifizierung organischer und metallorganischer Moleküle.

[0385] Absorptionsspektroskopie im nahen Infrarotbereich (NIR) ist die Messung der Wellenlänge und Intensität der Absorption von Licht im nahen Infrarotbereich durch eine Probe. Licht im nahen Infrarotbereich überspannt den Bereich  $800\text{ nm}$  bis  $2,5\text{ }\mu\text{m}$  ( $12.500$  bis  $4.000\text{ cm}^{-1}$ ) und ist energetisch ausreichend, um Obertöne und Kombinationen von Molekylvibrationen auf höhere Energieniveaus anzuregen. NIR-Spektroskopie wird typischerweise zur quantitativen Messung organischer funktioneller Gruppen verwendet, insbesondere O-H, N-H und C=O. Die Komponenten und die Konstruktion der NIR-Instrumentierung sind ähnlich zu UV-vis-Absorptionsspektrometern. Die Lichtquelle ist üblicherweise eine Wolframlampe und der Detektor ist üblicherweise ein PbS-Festkörperdetektor. Probenhalter können Quarz oder Glas sein und typische Lösungsmittel sind  $\text{CCl}_4$  und  $\text{CS}_2$ . Die bequeme Instrumentierung der NIR-Spektroskopie macht sie für on-line-Überwachung und Prozeßkontrolle geeignet.

[0386] Absorptionsspektroskopie im ultravioletten und sichtbaren Bereich (UV-vis) ist die Messung der Wellenlänge und Intensität der Absorption von Licht im nahen ultravioletten und sichtbaren Bereich durch eine Probe. Die Absorption im Vakuum-UV tritt bei  $100\text{--}200\text{ nm}$  ( $10^5\text{--}50.000\text{ cm}^{-1}$ ), Quarz-UV bei  $200\text{--}350\text{ nm}$  ( $50.000\text{--}28.570\text{ cm}^{-1}$ ) und im Sichtbaren bei  $350\text{--}800\text{ nm}$  ( $28.570\text{--}12.500\text{ cm}^{-1}$ ) auf und wird von dem Beer-Lambert-Bouguet-Gesetz beschrieben. Ultraviolettes und sichtbares Licht sind energetisch ausreichend, um äußere Elektronen in höhere Energieniveaus zu heben. UV-vis-Spektroskopie kann üblicherweise auf Moleküle und anorganische Ionen oder Komplexe in Lösung angewendet werden. Die UV-vis-Spektren sind durch die breiten Merkmale der Spektren begrenzt. Die Lichtquelle ist üblicherweise eine Wasserstoff- oder Deuteriumlampe für UV-Messungen und eine Wolframlampe für Messungen im Sichtbaren. Die Wellenlängen dieser kontinuierlichen Lichtquellen werden mit einem Wellenlängenseparator ausgewählt, wie etwa einem Prisma oder Gittermonochromator. Spektren werden durch Abtasten des Wellenlängenseparators erhalten und quantitative Messungen können aus einem Spektrum oder bei einer einzigen Wellenlänge durchgeführt werden.

[0387] Massenspektrometer nutzen den Unterschied im Masse/Ladungs-Verhältnis ( $m/z$ ) ionisierter Atome oder Moleküle, um sie voneinander zu trennen. Massenspektrometrie ist daher nützlich zur Quantifizierung von Atomen oder Molekülen und auch zur Bestimmung chemischer und struktureller Informationen über Moleküle. Moleküle haben deutliche Fragmentierungsmuster, die Strukturinformation liefern, um Verbindungen zu identifizieren. Die allgemeinen Betriebsweisen eines Massenspektrometers sind wie folgt. Gasphasen-Ionen werden erzeugt, die Ionen werden in Raum oder Zeit auf der Grundlage ihres Masse/Ladungs-Verhältnisses getrennt und die Menge an Ionen jeden Masse/Ladungs-Verhältnisses wird gemessen. Die Ionentrennkraft eines Massenspektrometers wird durch die Auflösung beschrieben, die definiert ist als  $R = m/\Delta m$ , wobei  $m$  die Ionenmasse und  $\Delta m$  der Unterschied in der Masse zwischen zwei auflösbaren Peaks in einem Massenspektrum ist. Ein Massenspektrometer mit einer Auflösung von  $1.000$  kann z. B. ein Ion mit einem  $m/z$  von  $100,0$  von einem Ion mit einem  $m/z$  von  $100,1$  auflösen.

[0388] Im allgemeinen besteht ein Massenspektrometer (MS) aus einer Ionenquelle, einem masseselektiven Analysator und einem Ionendetektor. Die Magnetsektor-, Quadrupol- und Flugzeit-Konstruktionen erfordern auch Extraktions- und Beschleunigungs-Ionenoptiken, um Ionen aus dem Quellenbereich in den Masseanalysatorbereich zu überführen. Die Details mehrerer Masseanalysatorkonstruktionen (für Magnetsektor-MS, Quadrupol-MS oder Flugzeit-MS) sind unten diskutiert. Einzelfokussierungs-Analysatoren für Magnetsektor-MS nutzen einen Teilchenstrahlweg von  $180$ ,  $90$  oder  $60$  Grad. Die verschiedenen Kräfte, die das Teilchen beeinflussen, trennen Ionen mit unterschiedlichen Masse/Ladungs-Verhältnissen. Bei Doppelfokussierungsanalysatoren wird ein elektrostatischer Analysator in diesem Instrumententyp hinzugefügt, um Teilchen mit Unterschieden in den kinetischen Energien zu trennen.

[0389] Ein Quadrupol-Massenfilter für Quadrupol-MS besteht aus vier Metallstäben, die parallel angeordnet sind. Die angelegten Spannungen beeinflussen die Bahn von Ionen, die sich entlang des Flugweges bewegen, der zwischen den vier Stäben zentriert ist. Für gegebene Gleich- und Wechselspannungen gehen nur Ionen eines bestimmten Masse/Ladungs-Verhältnisses durch den Quadrupolfilter hindurch und alle anderen Ionen werden aus ihrem ursprünglichen Weg geworfen. Ein Massenspektrum wird durch Überwachen der Ionen erhalten, die durch den Quadrupolfilter hindurchgehen, wenn die Spannungen an den Stäben variiert werden.

[0390] Ein Flugzeit-Massenspektrometer nutzt die Unterschiede in der Durchgangszeit durch eine "Driftregion", um Ionen unterschiedlicher Massen zu trennen. Er arbeitet in einem Pulsmodus, so daß Ionen in Pulsen erzeugt und/oder in Pulsen extrahiert werden müssen. Ein gepulstes elektrisches Feld beschleunigt alle Ionen in einer feldfreien Driftregion mit einer kinetischen Energie  $qV$ , wobei  $q$  die Ionenladung und  $V$  die angelegte Spannung ist. Da die kinetische Ionenenergie  $0,5\text{ mV}^2$  beträgt, haben leichtere Ionen eine höhere Geschwindigkeit als schwerere Ionen und erreichen den Detektor am Ende der Driftregion früher. Der Ausgang eines Ionendetektors wird auf einem Oszilloskop als eine Funktion der Zeit dargestellt, um das Massenspektrum zu

erzeugen.

[0391] Der Ionenbildungsprozeß ist der Startpunkt für massenspektrometrische Analysen. Chemische Ionisierung ist ein Verfahren, das ein Reagens-Ion einsetzt, um mit den Analytmolekülen (Markierungen) zu reagieren, um Ionen durch entweder einen Protonen- oder Hydridtransfer zu bilden. Die Reagens-Ionen werden erzeugt durch Einbringen eines großen Überschusses Methan (relativ zur Markierung) in eine Elektronenstoß(EI)-Ionenquelle. Elektronenzusammenstöße erzeugen  $\text{CH}_4^+$  und  $\text{CH}_3^+$ , die weiter mit Methan reagieren, um  $\text{CH}_5^+$  und  $\text{C}_2\text{H}_5^+$  zu bilden. Ein weiteres Verfahren, um Markierungen zu ionisieren, ist durch Plasma und Glühentladung. Plasma ist ein heißes, teilweise ionisiertes Gas, das Atome wirkungsvoll anregt und ionisiert. Eine Glühentladung ist ein Niederdruckplasma, das zwischen zwei Elektroden gehalten wird. Elektronenstoßionisation setzt einen Elektronenstrahl ein, üblicherweise von einem Wolframfaden erzeugt, um Gasphasen-Atome oder -Moleküle zu ionisieren. Ein Elektron aus dem Strahl schlägt ein Elektron aus Analyt-Atomen oder -Molekülen, um Ionen zu erzeugen. Elektrosprayionisierung verwendet eine sehr feine Nadel und eine Reihe von Abstreifern. Eine Probenlösung wird in die Quellenkammer gesprüht, um Tröpfchen zu bilden. Die Tröpfchen tragen Ladung, wenn sie aus der Kapillare austreten, und wenn das Lösungsmittel verdampft, verschwinden die Tröpfchen, wobei sie hochgeladene Analytmoleküle zurücklassen. ESI ist besonders nützlich für große biologische Moleküle, die schwierig zu verdampfen oder zu ionisieren sind. Bombardierung mit schnellen Atomen (FAB) verwendet einen Hochenergiestrahl neutraler Atome, typischerweise Xe oder Ar, der auf eine feste Probe trifft, wodurch Desorption und Ionisierung bewirkt wird. Es wird für große biologische Moleküle verwendet, die schwierig in die Gasphase zu bringen sind. FAB bewirkt geringe Fragmentierung und gibt üblicherweise einen großen Molekülionenpeak, was sie nützlich zur Bestimmung der relativen Molekülmasse macht. Der Atomstrahl wird durch Beschleunigung von Ionen aus einer Ionenquelle durch eine Ladungsaustauschzelle erzeugt. Ionen nehmen in Kollisionen mit neutralen Atomen ein Elektron auf, um einen Strahl aus Hochenergie-Atomen zu bilden. Laserionisierung (LIMS) ist ein Verfahren, bei dem ein Laserpuls Material von der Oberfläche einer Probe ablöst und ein Mikroplasma schafft, das einen Teil der Probenbestandteile ionisiert. MatriX-unterstützte Laserdesorptionsionisation (MALDI) ist ein LIMS-Verfahren zum Verdampfen und Ionisieren großer biologischer Moleküle, wie etwa Proteine oder DNA-Fragmente. Die biologischen Moleküle werden in einer festen Matrix, wie etwa Nicotinsäure, dispergiert. Ein UV-Laserpuls trägt Material von der Matrix ab, was einen Teil der großen Moleküle in einer ionisierten Form in die Gasphase bringt, so daß sie in einen Massenspektrometer extrahiert werden können. Plasmadesorptionsionisation (PD) verwendet den Zerfall von <sup>252</sup>Cf, was zwei Fissionsfragmente erzeugt, die sich in entgegengesetzte Richtungen bewegen. Ein Fragment trifft auf die Probe, wodurch 1–10 Analyt-Ionen herausgeschlagen werden. Das andere Fragment trifft auf einen Detektor und triggert den Start der Datenaufnahme. Dieses Ionisationsverfahren ist besonders nützlich für große biologische Moleküle. Resonanzionisation (RIMS) ist ein Verfahren, bei dem ein oder mehrere Laserstrahlen in Resonanz auf Übergänge eines Gasphasen-Atoms oder -Moleküls eingestellt werden, um es schrittweise über sein Ionisationspotential zu heben, um ein Ion zu erzeugen. Sekundärionisation (SIMS) verwendet einen Ionenstrahl, wie etwa  $^3\text{He}^+$ ,  $^{16}\text{O}^+$  oder  $^{40}\text{Ar}^+$ , wird auf die Oberfläche einer Probe fokussiert und zerstäubt Material in die Gasphase hinein. Funkenquelle ist ein Verfahren, das Analyte in feste Proben durch Pulsen eines elektrischen Stroms über zwei Elektroden ionisiert.

[0392] Eine Markierung kann vor, während oder nach der Abspaltung vom Molekül, an die sie gebunden ist, mit Ladung versehen werden. Ionisationsverfahren auf der Basis von Ionen-"Desorption", der direkten Bildung oder Emission von Ionen von festen oder flüssigen Oberflächen haben zunehmende Anwendung auf nicht-flüchtige und thermisch labile Verbindungen ermöglicht. Diese Verfahren eliminieren die Notwendigkeit der Verflüchtigung neutraler Moleküle vor der Ionisation und minimieren im allgemeinen die thermische Zersetzung der Molekülspezies. Diese Verfahren schließen Felddesorption (Becky, Principles of Field Ionization and Field Desorption Mass Spectrometry, Pergamon, Oxford, 1977), Plasmadesorption (Sundqvist und MacFarlane, Mass Spectrom. Rev. 4: 421, 1985), Laserdesorption (Karas und Hillenkamp, Anal. Chem. 60: 2299, 1988; Karas et al., Angew. Chem. 101: 805, 1989), Bombardierung mit schnellen Teilchen (z. B. Bombardierung mit schnellen Atomen, FAB, und Sekundärionen-Massenspektrometrie, SIMS, Barher et al., Anal. Chem. 54: 645A, 1982), und Thermospray(TS)-Ionisation (Vestal, Mass Spectrom. Rev. 2: 447, 1983) ein. Thermospray wird in breitem Umfang für die on-line-Kombination mit Flüssigkeitschromatographie angewendet. Die FAB-Verfahren mit kontinuierlichem Durchfluß (Caprioli et al., Anal. Chem. 58: 2949, 1986) haben ebenfalls signifikantes Potential gezeigt. Eine vollständigere Auflistung von Ionisations/Massenspektrometrie-Kombinationen ist Ionenfallen-Massenspektrometrie, Elektrosprayionisations-Massenspektrometrie, Ionenspray-Massenspektrometrie, Flüssigionisations-Massenspektrometrie, Massenspektrometrie mit Ionisation bei atmosphärischem Druck, Elektronenionisations-Massenspektrometrie, Massenspektrometrie mit Ionisation durch Bombardierung mit metastabilen Atomen, Massenspektrometrie mit Ionisation durch Bombardierung mit schnellen Atomen, MALDI-Massenspektrometrie, Photoionisations-Flugzeit-Massenspektrometrie, Lasertröpfchen-Massenspektrometrie, MALDI-TOF-Massenspektrometrie, APCI-Massenspektrometrie, Nanospray-Massenspektrometrie, Nebelsprayionisations-Massenspektrometrie, Massenspektrometrie mit chemischer Ionisation, Resonanzionisations-Massenspektrometrie, Sekundärionisations-Massenspektrometrie,

## Thermospray-Massenspektrometrie.

[0393] Die Ionisationsverfahren, die für nicht-flüchtige biologische Verbindungen geeignet sind, haben überlappende Anwendbarkeitsbereiche. Ionisationswirkungsgrade sind in hohem Maße abhängig von Matrixzusammensetzung und Verbindungstyp. Gegenwärtig verfügbare Ergebnisse weisen darauf hin, daß die obere Molekülmasse für TS etwa 8.000 Daltons beträgt (Jones und Krolik, Rapid Comm. Mass. Spektrom. 1: 67, 1987). Da TS hauptsächlich mit Quadrapol-Massenspektrometern durchgeführt wird, leidet die Empfindlichkeit typischerweise disproportional bei höheren Masse/Ladungs-Verhältnissen ( $m/z$ ). Flugzeit(TOF)-Massenspektrometer sind kommerziell erhältlich und besitzen den Vorteil, daß der  $m/z$ -Bereich nur durch den Detektorwirkungsgrad begrenzt ist. Vor kurzem sind zwei zusätzliche Ionisationsverfahren eingeführt worden. Diese zwei Verfahren werden nun als Matrix-unterstützte Laserdesorption (MALDI, Karas und Hillenkamp, Anal. Chem. 60: 2299, 1988; Karas et al., Angew. Chem. 101: 805, 1989) und Elektrosprayionisation (ESI) bezeichnet. Beide Methodiken haben einen sehr hohen Ionisationswirkungsgrad (d. h. sehr hohes [erzeugte Molekülionen]/[verbrauchte Moleküle]). Die Empfindlichkeit, die das letztendliche Potential der Technik definiert, ist abhängig von der Probengröße, der Menge an Ionen, der Durchflußgeschwindigkeit, dem Nachweisgrad und dem tatsächlichen Ionisationswirkungsgrad.

[0394] Elektrospray-MS beruht auf einer Idee, die als erstes in den 1960igern vorgeschlagen wurde (Dole et al., J. Chem. Phys. 49: 2240, 1968). Elektrosprayionisation (ESI) ist ein Mittel, um geladene Moleküle für die Analyse durch Massenspektroskopie zu erzeugen. Kurz gesagt erzeugt Elektrosprayionisation hochgeladene Tröpfchen durch Vernebeln von Flüssigkeiten in einem stark elektrostatischen Feld. Die hochgeladenen Tröpfchen, die im allgemeinen in einem Trockenbadgas bei atmosphärischem Druck gebildet werden, schrumpfen durch Verdampfung von neutralem Lösungsmittel, bis die Ladungsabstoßung die kohäsiven Kräfte überwindet, was zu einer "Coulomb-Explosion" führt. Der genaue Mechanismus der Ionisation wird kontrovers diskutiert und mehrere Gruppen haben Hypothesen vorgelegt (Blades et al., Anal. Chem. 63: 2109–14, 1991; Kobarle et al., Anal. Chem. 65: A972–86, 1993; Fenn, J. Am. Soc. Mass. Spektrom. 4: 524–35, 1993). Ungeachtet des letztendlichen Ionenbildungsprozesses erzeugt ESI geladene Moleküle aus Lösung unter milden Bedingungen.

[0395] Die Fähigkeit, nützliche Massenspektrumdaten von geringen Mengen eines organischen Moleküls zu erhalten, beruht auf der effizienten Erzeugung von Ionen. Der Ionisationswirkungsgrad für ESI steht im Zusammenhang mit dem Ausmaß der positiven Ladung, die mit dem Molekül assoziiert ist. Die Verbesserung der Ionisation auf experimentellem Wege hat üblicherweise die Verwendung saurer Bedingungen mit sich gebracht. Ein weiteres Verfahren, um die Ionisation zu verbessern, ist gewesen, quartäre Amine zu verwenden, wenn möglich (siehe Aebersold et al., Protein Science 1: 494–503, 1992; Smith et al., Anal. Chem. 60: 436–41, 1988).

[0396] Elektrosprayionisation wird detaillierter wie folgt beschrieben. Elektrospray-Ionenerzeugung erfordert zwei Schritte: Dispersion hochgeladener Tröpfchen bei nahezu atmosphärischem Druck, gefolgt von Zuständen, um Verdampfung zu induzieren. Eine Lösung von Analyt-Molekülen wird durch eine Nadel geleitet, die bei hohem elektrischen Potential gehalten wird. Am Ende der Nadel dispergiert die Lösung in einen Nebel von kleinen hochgeladenen Tröpfchen hinein, die die Analyt-Moleküle enthalten. Die kleinen Tröpfchen verdampfen schnell, und mit einem Verfahren der Felddesorption oder Restverdampfung werden protonierte Proteinmoleküle in die Gasphase freigesetzt. Ein Elektrospray wird im allgemeinen durch Anlegen eines hohen elektrischen Feldes an einen geringen Flüssigkeitsstrom (im allgemeinen 1–10  $\mu\text{l}/\text{min}$ ) aus einer Kapillarröhre erzeugt. Eine Potentialdifferenz von 3–6 kV wird typischerweise zwischen der Kapillare und der Gegenelektrode angelegt, die 0,2 bis 2 cm entfernt angeordnet ist (wo Ionen, geladene Cluster und sogar geladene Tröpfchen, in Abhängigkeit vom Ausmaß der Desolvatisierung, vom MS durch eine kleine Öffnung gesammelt werden können). Das elektrische Feld führt zu einer Ladungsakkumulation auf der Flüssigkeitsoberfläche des Kapillarendes; somit sind die Flüssigkeitsdurchflußgeschwindigkeit, der spezifische Widerstand und die Oberflächenspannung wichtige Faktoren bei der Tröpfchenenerzeugung. Das hohe elektrische Feld führt zum Aufbrechen der Flüssigkeitsoberfläche und zur Bildung hochgeladener Flüssigkeitströpfchen. Positiv oder negativ geladene Tröpfchen können in Abhängigkeit von der Kapillarspannung erzeugt werden. Der negative Ionenmodus erfordert das Vorhandensein eines Elektronenfängers, wie etwa Sauerstoff, um elektrische Entladung zu hemmen.

[0397] Ein breiter Bereich von Flüssigkeiten kann elektrostatisch in ein Vakuum gesprüht werden, oder mit Hilfe eines Vernebelungsmittels. Die Verwendung von ausschließlich elektrischen Feldern zur Vernebelung führt zu einigen praktischen Beschränkungen des Bereichs der Flüssigkeitsleitfähigkeit und dielektrischen Konstante. Lösungsleitfähigkeit von weniger als  $10^{-5}$  Ohm ist bei Raumtemperatur für ein stabiles Elektrospray bei brauchbaren Flüssigkeitsdurchflußgeschwindigkeiten erforderlich, was einer wäßrigen Elektrolytlösung von  $< 10^{-4}$  M entspricht. Bei dem Modus, der sich für ESI-MS als am nützlichsten erwiesen hat, führt eine geeignete Flüssigkeitsdurchflußgeschwindigkeit zu Dispersion der Flüssigkeit als ein feiner Nebel. In kurzer Entfernung von der Kapillare ist der Tröpfchendurchmesser oft ziemlich gleichförmig und in der Größenordnung von 1  $\mu\text{m}$ . Von besonderem Interesse ist, daß der Elektrospray-Ionenstrom für höhere Flüssigkeitsdurchfluß-

geschwindigkeiten nur insgesamt leicht ansteigt. Es gibt Hinweise, daß Erwärmung für die Manipulation des Elektrosprays nützlich ist. Leichte Erwärmung ermöglicht zum Beispiel, daß wäßrige Lösungen leicht in Elektrospray überführt werden können, vermutlich aufgrund der verringerten Viskosität und Oberflächenspannung. Sowohl thermisch unterstützte als auch Gasvernebelungs-unterstützte Elektrosprays ermöglichen, daß höhere Flüssigkeitsdurchflußraten verwendet werden können, senken aber das Ausmaß der Tröpfchenladung. Die Bildung von Molekülonen erfordert Bedingungen, die eine Verdampfung der anfänglichen Tröpfchenpopulation bewirken. Dies kann bei höheren Drücken durch einen Trockengasstrom bei mäßigen Temperaturen ( $< 60^{\circ}\text{C}$ ), durch Erhitzen während des Transports durch die Schnittstelle und (insbesondere im Falle von Ioneneinfangverfahren) durch energetische Kollisionen bei relativ niedrigem Druck durchgeführt werden.

[0398] Obgleich die detaillierten Prozesse, die ESI zugrundeliegen, unsicher bleiben, scheinen die sehr kleinen Tröpfchen, die durch ESI erzeugt werden, zu ermöglichen, daß nahezu jede Spezies, die in Lösung eine Nettoladung trägt, nach Verdampfung von Restlösungsmittel in die Gasphase überführt werden kann. Massenspektrometrischer Nachweis erfordert dann, daß Ionen nach Desolvatisierung einen verfolgbaren  $m/z$ -Bereich haben ( $< 4.000$  Daltons für Quadrupol-Instrumente) sowie mit ausreichendem Wirkungsgrad erzeugt und überführt werden können. Der breite Bereich von gelösten Stoffen, für den bereits festgestellt wurde, daß ESI-MS dafür geeignet ist, und das Fehlen einer substantiellen Abhängigkeit des Ionisationswirkungsgrades von der relativen Molekülmasse weist auf ein in hohem Maße nicht-diskriminierendes und breit anwendbares Ionisationsverfahren hin.

[0399] Die Elektrospray-Ionen-"Quelle" arbeitet bei nahezu atmosphärischem Druck. Die Elektrospray-"Quelle" ist typischerweise eine Metall- oder Glaskapillare, die ein Verfahren zur elektrischen Vorspannung der flüssigen Lösung relativ zu einer Gegenelektrode umfaßt. Lösungen, typischerweise Wasser-Methanol-Mischungen, die den Analyten und oft weitere Additive, wie etwa Essigsäure, enthalten, strömen zum Kapillarenden. Eine ESI-Quelle ist beschrieben worden (Smith et al., Anal. Chem. 62: 885, 1990), die im wesentlichen jedes Lösungsmittelsystem aufnehmen kann. Typische Durchflußgeschwindigkeiten für ESI sind 1 bis  $10\ \mu\text{l}/\text{min}$ . Das prinzipielle Erfordernis einer ESI-MS-Schnittstelle ist, Ionen aus dem Hochdruckbereich so wirkungsvoll wie möglich in das MS zu sampeln und zu transportieren.

[0400] Der Wirkungsgrad von ESI kann sehr hoch sein, was die Grundlage für extrem empfindliche Messungen liefert, was für die hierin beschriebene Erfindung nützlich ist. Die augenblickliche Instrumentenleistung kann einen Gesamtionenstrom am Detektor von etwa  $2 \times 10^{-12}\ \text{A}$  oder etwa  $10^7$  Zählpulse/s für einfach geladene Spezies liefern. Auf der Basis der Instrumentenleistung werden so niedrige Konzentrationen wie  $10^{-10}\ \text{M}$  oder etwa  $10^{-16}\ \text{mol/s}$  einer einfach geladenen Spezies einen nachweisbaren Ionenstrom (etwa  $10$  Zählpulse/s) ergeben, wenn der Analyt vollständig ionisiert ist. Nachweisgrenzen im unteren Attomol-Bereich sind z. B. für quartäre Ammonium-Ionen unter Verwendung einer ESI-Schnittstelle mit Kapillarzonenelektrophorese erhalten worden (Smith et al., Anal. Chem. 59: 1230, 1988). Für eine Verbindung mit einer relativen Molekülmasse von 1.000 ist die mittlere Zahl von Ladungen 1, die mittlere Zahl von Ladungszuständen 1, die Peakbreite ( $m/z$ ) 1 und die maximale Intensität (Ionen/s)  $1 \times 10^{12}$ .

[0401] Eine bemerkenswert kleine Probe wird beim Erstellen eines ESI-Massenspektrums tatsächlich verbraucht (Smith et al., Anal. Chem. 60: 1948, 1988). Substantielle Gewinne könnten auch durch die Verwendung von Array-Detektoren mit Sektorinstrumenten erzielt werden, was gleichzeitigen Nachweis von Teilen des Spektrums ermöglicht. Da gegenwärtig nur etwa  $10^{-5}$  aller Ionen, die von ESI gebildet werden, nachgewiesen werden, kann eine Fokussierung auf die Faktoren, die die Instrumentenleistung begrenzen, eine Grundlage für verbesserte Empfindlichkeit liefern. Es wird den Fachleuten deutlich sein, daß die vorliegende Erfindung Verbesserungen für Ionisations- und Nachweismethodiken überdenkt und anpaßt.

[0402] Vorzugsweise wird zwischen die Trennvorrichtung (z. B. Gel) und den Detektor (z. B. Massenspektrometer) eine Schnittstelle gesetzt. Die Schnittstelle hat vorzugsweise die folgenden Eigenschaften: (1) die Fähigkeit, die DNA-Fragmente in diskreten Zeitintervallen zu sammeln, (2) die DNA-Fragmente zu konzentrieren, (3) die DNA-Fragmente aus den Elektrophoresepuffern und dem Milieu zu entfernen, (4) die Markierung vom DNA-Fragment abzuspalten, (5) die Markierung vom DNA-Fragment zu trennen, (6) das DNA-Fragment zu verwerfen, (7) die Markierung in eine flüchtige Lösung zu geben, (8) die Markierung zu verflüchtigen und zu ionisieren und (9) die Markierung in eine Elektrosprayvorrichtung einzubringen oder dorthin zu transportieren, die die Markierung in das Massenspektrometer einführt.

[0403] Die Schnittstelle hat auch die Fähigkeit, DNA-Fragmente zu "sammeln", wenn sie aus dem Boden eines Gels eluieren. Das Gel kann aus einem Slabgel, einem Röhrengel, einer Kapillare, etc. bestehen. Die DNA-Fragmente können mit verschiedenen Verfahren gesammelt werden. Das erste Verfahren ist dasjenige der Verwendung eines elektrischen Feldes, wobei DNA-Fragmente auf oder nahe einer Elektrode gesammelt werden. Ein zweites Verfahren ist dasjenige, bei dem die DNA-Fragmente durch Leiten eines Flüssigkeitsstroms am Boden eines Gels vorbei gesammelt werden. Aspekte beider Verfahren können kombiniert werden, wobei DNA in einen fließenden Strom hinein gesammelt wird, der später unter Verwendung eines elektrischen Feldes konzentriert werden kann. Das Endergebnis ist, daß DNA-Fragmente aus dem Milieu entfernt werden, in dem das Trennverfahren durchgeführt wurde. Das heißt, daß DNA-Fragmente durch Verwendung eines

elektrischen Feldes aus einem Lösungstyp in einen anderen "gezogen" werden können.

[0404] Wenn die DNA-Fragmente erst einmal in der geeigneten Lösung sind (kompatibel mit Elektrospray und Massenspektrometrie), kann die Markierung vom DNA-Fragment abgespalten werden. Das DNA-Fragment (oder Reste desselben) kann von der Markierung durch das Anlegen eines elektrischen Feldes getrennt werden (vorzugsweise hat die Markierung die entgegengesetzte Ladung zu derjenigen der DNA-Markierung). Die Markierung wird anschließend durch die Verwendung eines elektrischen Feldes oder einer strömenden Flüssigkeit in die Elektrosprayvorrichtung eingebracht.

[0405] Die Nachweis-Vorrichtung ist bevorzugterweise ein Massenspektrometer. Weil Massenspektrometer den Unterschied in den Masse-zu-Ladungsverhältnis ( $m/e$ ) von ionisierten Spezies verwenden, um spezifisch Moleküle zu identifizieren, ist diese Nachweisteknik zur Quantifizierung von kleinen Molekülen brauchbar und auch zur Bestimmung von chemischen und strukturellen Informationen über Moleküle. Einige Moleküle weisen distinkte Fragmentierungsmuster auf, die Information zur Verfügung stellen können, um strukturelle Komponenten zu identifizieren. Jedoch, für die hier beschriebene Verwendung wird das MSD im wesentlichen als ein "Any-Detektor" für den Nachweis, die Messung und die Quantifizierung von Markern von bekannten Molekulargewicht verwendet. Daher kann ein Massenspektrometer so ähnlich verwendet werden, wie Wissenschaftler im Augenblick einen Dioden-Array-Detektor in einen UVNIS-Spektrometer verwenden, um kleine Moleküle mit bekannten Extinktionskoeffizienten zu messen. In der hier beschriebenen Anwendung werden die Marker dazu verwendet, um die Anwesenheit oder Abwesenheit von spezifischen Nukleinsäuresequenzen zu identifizieren und die Probenidentifikation zu kartieren.

[0406] Ein Quadrupol-Massendetektor besteht aus vier parallelen Metallstäben. Zwei gegenüberliegende Stäbe weisen ein angelegtes Potential von  $(U-V\cos(wt))$  auf und die anderen zwei Stäbe weisen ein Potential von  $-(U-V\cos(wt))$  auf, worin  $U$  eine DC-Spannung und  $V\cos(wt)$  eine AC-Spannung ist. Die angelegten Spannungen beeinträchtigen die Trajektion von Ionen, die entlang der Flugstrecke wandern, die zwischen den vier Stäben zentriert ist. Für gegebene DC- und AC-Spannungen passieren nur die Ionen eines bestimmten Masse-zu-Ladungsverhältnisses durch den Quadrupol-Filter und alle anderen Ionen werden aus ihrer Ausgangsbahn beworfen. Ein Massenspektrum wird durch Überwachen der Ionen erhalten, die durch den Quadrupol-Filter passieren, während die Spannungen an den Stäben variiert werden. Die Ionentrennfähigkeit eines Massenspektrometers wird durch die Auflösung beschrieben, die die Differenz in der Masse zwischen zwei auflösbaren Peaks in einem Massenspektrum ist. Gemeint ist, ein Massenspektrometer mit einer Auflösung von 1000 kann ein Ion mit einer  $m/e$  von 100,0 von einem Ion mit einer  $m/e$  von 100,1 auflösen. Der allgemeine Betrieb eines Massenspektrometers ist es, zuerst Gas-Phase-Ionen zu erzeugen, zweitens, die Ionen im Raum oder in der Zeit basierend auf ihren Masse-zu-Ladungsverhältnis zu trennen und drittens, die Menge von Ionen von jedem Masse-zu-Ladungsverhältnis zu messen.

[0407] Das Massenspektrometer ist ideal als ein Spektrometer für Anwendungen in Genomics geeignet, da es die gleichzeitige Messung von hunderten von Markern erlaubt. Die augenblickliche Zahl von genomischen Anwendungen (Sequenzierung, Kartierung, Genotypisierung) verwendet an Markern, ist ungefähr 4, was sich aus den überlappenden Emissionsspektrum von fluoreszenten Markern ergibt, die zwischen 300 nm und 700 nm plaziert werden können. Im Unterschied dazu können mit augenblicklich verwendeten Quadrupol-Instrumenten (sowie dem Micromass MS, dem Hewlett Packard LC/MSD 1100, den PE Sciex API 165 LC/MS oder dem Finnigan Navigator) ungefähr 400 Marker in den Spektren von 50–3000 amu plaziert werden. Die MS-Instrumente weisen mindestens 0,1 amu auf Lösung auf. Das neue Meßsystem, das wir beschreiben (d. h. markierte Biomoleküle) kann in Verbindung mit fast allen kommerziellen Massenspektrometern und HPLC-Systemen mit geringen Modifikationen verwendet werden. Idealerweise wird ein Software-Paket mit der "Driver"-Software beigelegt werden, um die molekularbiologischen, genetischen oder genomischen Anwendungen zu erarbeiten. Geeignete Software-Pakete sind beschrieben in z. B. U.S. Provisional Patent Application Nr. 60/053,429, angemeldet am 22. Juli 1997.

[0408] Atmosphärischer Druck chemischer Ionisierung (APCI) kann auf eine Vielzahl von Markertypen angewendet werden und wird im allgemeinen dazu verwendet, die Häufigkeit des molekularen Ions (Moleküle Formen seltener Addukte während APCI) zu verstärken. Der brauchbare Molekulargewichtsbereich von APCI ist 50 bis 3000 amu (im allgemeinen). Die Genauigkeit der Messung bei niedriger Auflösungsfähigkeit ist 0,1 amu und in dem hochauflösenden Modus, 5 ppm. Die APCI verwendet ein Reagenz-Ion, um mit den Analytmolekülen zu reagieren, um Ionen durch entweder einen Protonen- oder Hydridtransfer zu bilden. Im Augenblick werden nur ungefähr  $10^{-4}$  bis  $10^{-5}$  von allen durch APCI gebildeten Ionen nachgewiesen. Dieser wichtige Parameter (der die Instrumenten-Performance begrenzt) kann eine Basis für erhöhte Sensitivität zur Verfügung stellen. Diese Ionisierungstechnik ist ein kontinuierliches Verfahren, daß zur Verwendung als eine Schnittstelle mit HPLC oder Kapillar-Elektrophorese geeignet ist.

[0409] Es gibt alternative Formen von Ionisation, die für die CMST-Technologie angewendet werden können. Die Elektrospray-Ionisation (ESI) ermöglicht die Produktion von Molekularionen, die direkt aus Proben in Lösung. Die Elektrospray-Ionisation ist ein Verfahren, das mit einzelnen Quadrupol-Instrumenten kompatibel ist. Sehr wenig Probe wird während des Erhalts eines ESI-Massenspektrums (Smith et al., Anal. Chem. 60: 1948,

1988) gebraucht, jedoch bleibt die Gesamteffizienz der Ioneneinführung in das MS-Vakuumsystem relativ ineffizient. Die ESI leidet auch an den Effekten von denaturierenden Mitteln und Detergentien, die den Ionisationsschritt nachteilig beeinflussen. Sie kann auch für Biopolymere kleinen und großen Molekulargewichts (Peptide, Proteine, Kohlenhydrate und DNA-Fragmente) und Lipide verwendet werden. Anders als MALDI, die gepulst ist, ist ESI ein kontinuierliches Ionisationsverfahren. Mit ESI als der Ionisationsmethode werden mehrfach geladene Ionen normalerweise hergestellt (Moleküle neigen mehr zur Adduktbildung).

[0410] Ein alternatives Ionisationsverfahren ist die Matrix-unterstützte Laser-Desorption (MALDI), die dazu verwendet werden kann, um das Molekulargewicht von hier beschriebenen Markern zu bestimmen (Peptide, Proteine, Oligonukleotide und andere Verbindungen von biologischen Ursprung, sowie von kleinen synthetischen Polymeren können auch gemessen werden). Die Menge von erforderlicher Probe ist sehr gering (pmol oder weniger). Die Analyse kann auf eine lineare Weise (hohe Masse, niedriger Auflösung) bis zu einem Molekulargewicht von  $m/z$  300.000 (in seltenen Fällen) oder Reflektronmodus (geringere Masse, höhere Auflösung) bis zu einem Molekulargewicht von 10.000 durchgeführt werden.

[0411] APCI und ESI sollten im allgemeinen als ein Komplement zu MALDI betrachtet werden. MALDI weist den schweren Nachteil auf, ein nicht-quantifizierbares Ionisationsverfahren zu sein. Die Elektrospray-Ionisation wurde auf HP-Quadrapol, den Finnigan LCQ, den Finnigan TSQ 7000 (Medicine), den PE/Sciex-Instrument und dem Micromass-Instrument installiert. Sowohl APCI und ESI weisen den Vorteil über MALDI auf, daß sie in der Lage sind, Echtzeitmessungen von Markern während jeder Art von Abtrenntechnologie (d. h. HPLC und Elektrophorese) durchzuführen.

[0412] Fluoreszenz-Markierungen können am direktesten durch ihre Absorptions- und Fluoreszenzemissionswellenlängen und -intensitäten identifiziert und quantifiziert werden.

[0413] Während ein herkömmlicher Spektrofluorometer extrem flexibel ist, wobei er kontinuierliche Bereiche für die Anregungs- und Emissionswellenlängen liefert ( $I_{EX}$ ,  $I_{S1}$ ,  $I_{S2}$ ), können spezialisiertere Instrumente, wie etwa Durchflußcytometer und Laserabtastmikroskope, Sonden erfordern, die bei einer einzigen fixierten Wellenlänge anregbar sind. In heutigen Instrumenten ist dies üblicherweise die 488-nm-Linie des Argonlasers.

[0414] Die Fluoreszenzintensität pro Sondenmolekül ist proportional zum Produkt von E und QY. Der Bereich dieser Parameter unter Fluorophoren von gegenwärtiger praktischer Bedeutung ist ungefähr 10.000 bis 100.000  $\text{cm}^{-1}\text{M}^{-1}$  für e und 0,1 bis 1,0 für QY. Wenn die Absorption durch Hochintensitätsbeleuchtung zur Sättigung getrieben wird, wird die irreversible Zerstörung des angeregten Fluorophors (Photobleichung) der Faktor, der die Fluoreszenznachweisbarkeit begrenzt. Die praktische Auswirkung der Photobleichung hängt von der jeweiligen Fluoreszenznachweisteknik ab.

[0415] Es wird einem Fachmann deutlich sein, daß eine Vorrichtung (eine Schnittstelle) zwischen die Trenn- und Nachweisschritte geschaltet werden kann, um den kontinuierlichen Betrieb von Größentrennung und Markierungsnachweis (in Echtzeit) zu ermöglichen. Dies vereinheitlicht die Trennmethode und -instrumentierung mit der Nachweismethode und -instrumentierung, unter Bildung einer einzigen Vorrichtung. Eine Schnittstelle ist z. B. zwischen einer Trenntechnik und den Nachweis durch Massenspektrometrie oder potentiostatischer Amperometrie geschaltet.

[0416] Die Funktion der Schnittstelle ist primär die Freisetzung der (z. B. Massenspektrometrie)-Markierung vom Analyten. Es gibt mehrere repräsentative Implementierungen der Schnittstelle. Die Konstruktion der Schnittstelle ist abhängig von der Auswahl der spaltbaren Linker. Im Falle von licht- oder photospaltbaren Linkern ist eine Energie- oder Photonenquelle erforderlich. Im Falle eines säurelabilen Linkers, eines basenlabilen Linkers oder eines Disulfid-Linkers ist die Zugabe von Reagens in der Schnittstelle erforderlich. Im Falle wärmeablabiler Linker ist eine Energiewärmequelle erforderlich. Enzymzugabe ist erforderlich für einen enzymempfindlichen Linker, wie etwa eine spezifische Protease und einen Peptid-Linker, einen Nuklease- und einen DNA- oder RNA-Linker, eine Glycosylase, HRP oder Phosphatase oder einen Linker, der nach Spaltung instabil ist (z. B. ähnlich zu Chemilumineszenzsubstraten). Andere Eigenschaften der Schnittstelle schließen minimale Bandenverbreiterung, Trennung der DNA von Markierungen vor Injektion in einen Massenspektrometer ein. Trenntechniken schließen diejenigen ein, die auf Elektrophoreseverfahren und -techniken, Affinitätstechniken, Größenretention (Dialyse), Filtration und dergleichen beruhen.

[0417] Es ist auch möglich, die Markierungen (oder das Nukleinsäure-Linker-Markierungs-Konstrukt) zu konzentrieren, elektrophoretisch einzufangen und dann in einen alternativen Reagensstrom freizusetzen, der mit dem bestimmten ausgewählten Ionisationsverfahrenstyp kompatibel ist. Die Schnittstelle kann auch in der Lage sein, die Markierungen (oder Nukleinsäure-Linker-Markierungs-Konstrukte) auf Mikroperlen einzufangen, die Perle(n) in eine Kammer einzuschließen und anschließend Laserdesorption/Verdampfung durchzuführen.

[0418] Es ist auch möglich, im Strom in einen alternativen Puffer hinein zu extrahieren (z. B. aus Kapillarelektrophoresepuffer in hydrophoben Puffer durch eine permeable Membran). In einigen Anwendungen kann es auch wünschenswert sein, Markierungen intermittierend in den Massenspektrometer einzuführen, was eine weitere Funktion der Schnittstelle umfassen würde. Eine weitere Funktion der Schnittstelle ist, Markierungen aus mehreren Säulen in einen Massenspektrometer zuzuführen, mit einem rotierenden Zeitslot für jede Säule.



Es ist auch möglich, Markierungen aus einer einzigen Säule in mehrere MS-Detektoren einzuführen, getrennt nach Zeit, jeden Satz von Markierungen für ein paar Millisekunden zu sammeln und dann einem Massenspektrometer zuzuführen.

[0419] Die folgende ist eine Liste repräsentativer Anbieter für Trenn- und Nachweisttechnologien, die in der vorliegenden Erfindung verwendet werden können. Hoefer Scientific Instruments (San Francisco, CA) stellt Elektrophoreseausrüstung (Two Step<sup>®</sup>, Poker Face<sup>®</sup>II) für Sequenzierungsanwendungen her. Pharmacia Biotech (Piscataway, NJ) stellt Elektrophoreseausrüstung für DNA-Trennungen und -sequenzierung her (Phast-System für PCR-SSCP-Analyse, MacroPhor System für DNA-Sequenzierung). Perkin Elmer/Applied Biosystems Division (ABI, Foster City, CA) stellt halbautomatisierte Sequenzer auf der Basis von Fluoreszenzfarbstoffen her (ABI373 und ABI377). Analytical Spectral Devices (Boulder, CO) stellt UV-Spektrometer her. Hitachi Instruments (Tokio, Japan) stellt Atomabsorptionsspektrometer, Fluoreszenzspektrometer, LC- und GC-Massenspektrometer, NMR-Spektrometer und UV-vis-Spektrometer her. PerSeptive Biosystems (Framingham, MA) stellt Massenspektrometer (Voyager<sup>®</sup> Elite) her. Bruker Instrumens, Inc. (Manning Park, MA) stellt FTIR-Spektrometer (Vector 22), FT-Raman-Spektrometer, Flugzeit-Massenspektrometer (Reflex II<sup>®</sup>, Ionenfallen-Massenspektrometer (Esquire<sup>®</sup>) und einen Maldi-Massenspektrometer her. Analytical Technology Inc. (ATI, Boston, MA) stellt Kapillargelelektrophoreseeinheiten, UV-Detektoren und Diodenarray-Detektoren her. Teledyne Electronic Technologies (Mountain View, CA) stellt einen Ionenfallen-Massenspektrometer her (3DQ Discovery<sup>®</sup> und den 3DQ Apogee<sup>®</sup>). Perkin Elmer/Applied Biosystems Division (Foster City, CA) stellt einen Sciex-Massenspektrometer (Dreifachquadrupol LC/MS/MS, den API 100/300) her, der mit Elektrospray kompatibel ist.

[0420] Hewlett-Packard (Santa Clara, CA) stellt massenselektive Detektoren (HP 5972A), MALDI-TOF-Massenspektrometer (HP G2025A), Diodenarray-Detektoren, CE-Einheiten, HPLC-Einheiten (HP 1090) sowie UV-Spektrometer her. Finnigan Corporation (San Jose, CA) stellt Massenspektrometer (Magnetsektor (MAT 95 S<sup>®</sup>), Quadrupol-Spektrometer (MAT 95 SQ<sup>®</sup>) und vier weitere verwandte Massenspektrometer) her. Rainin (Ameryville, CA) stellt HPLC-Instrumente her.

[0421] Die Verfahren und Zusammensetzungen, die hierin beschrieben sind, erlauben die Verwendung von abgespaltenen Markierungen, um als Karten für bestimmte Probenotypen und Nukleotididentitäten zu dienen. Zu Beginn jedes Sequenzierungsverfahrens wird ein bestimmter (ausgewählter) Primer einer bestimmten einzigartigen Markierung zugeordnet. Die Markierungen kartieren auf entweder einen Probenotyp, einen Didesoxyterminortyp (im Falle einer Sanger-Sequenzierungsreaktion) oder vorzugsweise beides. Genauer gesagt kartiert die Markierung auf einen Primertyp, der seinerseits auf einen Vektortyp kartiert, der seinerseits auf eine Probenidentität kartiert. Die Markierung kann auch auf einen Didesoxyterminortyp (ddTTP, ddCTP, ddGTP, ddATP) durch Bezugnahme, in welche Didesoxynukleotidreaktion der markierte Primer gegeben wird, kartieren. Die Sequenzierungsreaktion wird dann durchgeführt, und die resultierenden Fragmente werden sequentiell über die Zeit nach Größe getrennt.

[0422] Die Markierungen werden von den Fragmenten in einem Zeitrahmen abgespalten und in einem Zeitrahmen gemessen und aufgezeichnet. Die Sequenz wird durch Vergleich der Markierungskarte mit dem Zeitrahmen konstruiert. Das heißt, daß alle Markierungsidentitäten über die Zeit nach dem Bemessungsschritt aufgezeichnet und miteinander in einem Zeitrahmen in Beziehung gesetzt werden. Der Bemessungsschritt trennt die Nukleinsäurefragmente nach Inkrementen von einem Nukleotid, und die damit zusammenhängenden Markierungsidentitäten werden daher nach Inkrementen von einem Nukleotid getrennt. Durch Vorkenntnis des Didesoxyterminators, oder der Nukleotidkarte und des Probenotyps kann die Sequenz leicht auf lineare Art und Weise abgeleitet werden.

[0423] Ein genetisches Fingerprinting-System der vorliegenden Erfindung besteht im allgemeinen aus einer Probeneinführungsvorrichtung, einer Vorrichtung, um die markierten Proben von Interesse zu trennen, eine Aufteilverrichtung, um eine veränderliche Menge der Probe zu einem Fraktionssammler umzuleiten, eine Vorrichtung, um die Marker von den Proben von Interesse abzuspalten, eine Vorrichtung zum Nachweis des Markers und ein Softwareprogramm, um die gesammelten Daten zu analysieren und sie in einem Differential-Display-Modus wiederzugeben. Es wird dem Fachmann ersichtlich sein, wenn er im Besitz der vorliegenden Beschreibung ist, daß diese allgemeine Beschreibung viele Varianten jeder der angeführten Komponenten aufweisen kann. Wie am besten in **Fig. 15** gesehen, besteht ein beispielhaftes genetisches Fingerprinting-System **10** der vorliegenden Erfindung aus einer Probeneinführungsvorrichtung **12**, einer Trennvorrichtung **14**, die die Proben durch High-Performance-Liquid-Chromatographie (HPLC) trennt, eine Aufteilverrichtung **13**, einen Fraktionssammler **15**, eine Photospaltungsvorrichtung **16**, um die Marker von den Proben von Interesse abzuspalten, eine Nachweisvorrichtung **18**, die die Marker mittels eines elektrochemischen Detektors nachweist und eine Daten verarbeitende Vorrichtung **20**, mit einem Datenanalyse-Softwareprogramm, das die Ergebnisse aus der Nachweisvorrichtung analysiert. Jede Komponente wird im folgenden genauer diskutiert.

[0424] Die Probeneinführungsvorrichtung **12** nimmt automatisch einen gemessenen Aliquot **22** des PCR-Produkts, das in dem genetischen Fingerprinting-Verfahren erzeugt wurde und liefert sie über eine herkömmliche Leitung **24** zu der Trennvorrichtung **14** (im allgemeinen eine HPLC). Die Probeneinführungsvorrichtung **12** der ex-

emplarischen Ausführungsform besteht aus einem Temperatur kontrollierten Autosampler **26**, der Mikrotiterplatten aufnehmen kann. Der Autosampler **26** ist Temperatur kontrolliert, um die Integrität der erzeugten Nukleinsäureproben aufrechtzuerhalten und ist in der Lage, 25 µl oder weniger an Probe zu injizieren. Hersteller von diesem Typ von Probeneinführungsvorrichtung **12** sind, zum Beispiel, Gilson (Middleton, WI).

[0425] Die Probeneinführungsvorrichtung **12** ist operativ in Reihe mit der Auftrennvorrichtung **14** durch eine herkömmliche Leitung **24** verbunden. Die PCR-Produkte in den gemessenen Aliquot **22**, die in der Auftrennvorrichtung **14** aufgenommen werden, werden temporal durch High-Performance-Liquid-Chromatographie getrennt, um getrennte DNA-Fragmente zur Verfügung zu stellen. Die High-Performance-Liquid-Chromatographie kann eine isokratische, binäre oder quaternäre Pumpe(n) **27** aufweisen und kann von vielen Herstellern erworben werden (z. B. Hewlett Packard (Palo Alto, CA) HP 1100 oder 1090 Serie, Beckman Instruments Inc. (800-142-2345), Bioanalytical Systems, Inc. (800-845-4246), ESA, Inc. (508) 250-700), Perkin-Elmer Corp. (800-762-4000), Varian Instruments (800-926-3000), Waters Corp. (800-254-4752)).

[0426] Die Auftrennvorrichtung **14** schließt eine analytische HPLC-Säule **28** ein, die zur Verwendung zur Auftrennung der Oligonukleotide geeignet ist. Die Säule **28** ist eine analytische HPLC, zum Beispiel modifizierter nicht-poröses Polystyroldivinylbenzol-(2,2 µm Partikelgröße)-fester Träger, der innerhalb eines pH-Bereichs von 2–12, Drücken von bis zu 3000 psi und einem Temperaturbereich von 10 bis 70°C betrieben werden kann. Eine Temperatur-Kontrollvorrichtung (z. B. ein Säulenofen) (nicht gezeigt) kann dazu verwendet werden, um die Temperatur der Säule zu kontrollieren. Solche Temperatur-Kontrollvorrichtungen sind im Stand der Technik bekannt und können zum Beispiel von Rainin Instruments (Tochterunternehmen von Varian Instruments, Palo Alto, CA) erworben werden. Eine geeignete Säule **28** ist unter dem kommerziellen Namen DNASep<sup>®</sup> erhältlich und ist von Serasep (San Jose, CA) erhältlich. Andere geeignete analytische HPLC-Säulen sind von anderen Herstellern erhältlich (z. B. Hewlett Packard (Palo Alto, CA), Beckman Industries (Brea, CA), Waters Corp. (Milford, MA) und Supelco (Bellefonte, PA)).

[0427] Die Auftrennvorrichtung **14** in der gezeigten Ausführungsform enthält den Probenteiler **13** und der Probenteiler ist mit dem Flußstrom der Proben verbunden. Der Probenteiler **13** ist dazu angepaßt, um auf eine herkömmliche Weise variable Mengen von Probe zum Fraktionssammler **15** entweder für weitere Analyse oder Lagerung umzuleiten. Der Fraktionssammler **15** muß in der Lage sein, kleine Volumen aufzunehmen, eine Temperaturkontrolle zu niedrigen Temperaturen aufzuweisen und einstellbare Probenzeiten zu haben. Hersteller von In-line-Abteilern schließen Upchurch (Oak Harbor, WA) ein.

[0428] Der Fraktionssammler **15** ist an die HPLC/LC-Vorrichtung über eine Fluß-Aufteilungsleitung **29** angebracht. Die Fraktionssammler **15** erlauben die Sammlung von spezifischen Peaks, DNAs, RNAs und Nukleinsäurefragmenten oder Molekülen von Interesse in Röhrchen, Vertiefungen von Mikrotiterplatten oder Behältern. Zusätzlich können die Fraktionssammler **15** alles oder einen Teil eines Sets von Nukleinsäurefragmenten sammeln, die durch HPLC oder LC getrennt wurden. Hersteller von Fraktionssammlern schließen Gilson (Middleton; WI) und Isco (Lincoln, NE) ein. Die Verwendung eines Fraktionssammlers **15** in dieser Technologie stellt bemerkenswerte wesentliche Vorteile über die Gel-basierten Systeme zur Verfügung. Zum Beispiel ist es möglich, direkt Nukleinsäurefragmente zu klonieren, die durch die HPLC- oder LC-Verfahren erhalten wurden. Zusätzlich ist es möglich, Nukleinsäurefragmente durch PCR zu amplifizieren, die durch HPLC- oder LC-Verfahren erhalten wurden. Diese zwei Verfahren ermöglichen die schnelle Identifizierung von Nukleinsäurefragmenten von Interesse auf einem Sequenzniveau. Beide Verfahren sind mühsam und ineffektiv, wenn in Verbindung mit Gel-basierten Systemen verwendet.

[0429] In der dargestellten Ausführungsform ist der Fraktionssammler **15** eine einzelne Komponente vom genetischen Fingerprinting-System **10**. In einer alternativen Ausführungsform (nicht gezeigt) ist der Fraktionssammler **15** in der Probeneinführungsvorrichtung **12** integriert. Demzufolge richtet die Fluß-Teilungsleitung **32** die geteilte Probe von dem Probenteiler **13** zurück zu der Probeneinführungsvorrichtung **12**.

[0430] Ein Strom der getrennten DNA-Fragmente (z. B. Sequenzreaktionsprodukte) fließt durch eine herkömmliche Leitung **30** von der Trennvorrichtung **14** stromabwärts des Probenteilers **13** zu der Spaltungsvorrichtung **16**. Jedes der DNA-Fragmente ist mit einem einmaligen abspaltbaren (z. B. Photo-abspaltbaren) Marker markiert. Der Flußstrom von aufgetrennten DNA-Fragmenten passiert durch oder entlang der Spaltvorrichtung **16**, wo der Marker für den Nachweis (z. B. durch Massenspektrometrie oder mit einem elektrochemischen Detektor) entfernt wird. In der beispielhaften Ausführungsform ist die Spaltvorrichtung **16** eine photospaltende Einheit, so daß der Flußstrom der Probe gegenüber ausgewählter Lichtenergie und Wellenlänge ausgesetzt wird. In einer Ausführungsform tritt die Probe in die photospaltende Einheit **16** ein und wird gegenüber der ausgewählten Lichtquelle für eine ausgewählte Zeitdauer ausgesetzt. In einer alternativen Ausführungsform wird der Flußstrom der Probe angrenzend an die Lichtquelle entlang eines Flußpfades getragen, der ein ausreichendes Aussetzen gegenüber seiner Lichtenergie zur Verfügung stellt, um die Marker von den getrennten DNA-Fragmenten abzuspalten. Eine photospaltende Einheit ist von Supelco (Bellefonte, PA) erhältlich. Photospaltung kann bei vielen Wellenlängen mit einer Quecksilber/Xenon-Lichtbogenlampe durchgeführt werden. Die Wellenlängengenauigkeit ist ungefähr 2 nm mit einer Bandbreite von 10 nm. Der bestrahlte Bereich ist kreisförmig und typischerweise ein Bereich von 10–100 Quadratzentimetern. In alternativen Ausführungsfor-

men können andere Spaltvorrichtungen, die durch Säure, Base, Oxidation, Reduktion, Fluorid, Thiolaustausch, Photolyse oder enzymatische Bedingungen spalten, dazu verwendet werden, um die Marker von den aufgetrennten DNA-Fragmenten zu entfernen.

[0431] Nachdem die Spaltvorrichtung **16** die Marker von den aufgetrennten DNA-Fragmenten abspaltet, fließen die Marker durch eine herkömmliche Röhre **32** zur Nachweisvorrichtung **18** zum Nachweis jedes Markers. Der Nachweis der Marker kann auf dem Unterschied im elektrochemischen Potential zwischen jedem der Marker basiert sein, die dazu verwendet wurden, jede Art von DNA zu markieren, die in dem PCR-Schritt erzeugt wurde. Der elektrochemische Detektor **18** kann auf entweder coulometrischen oder amperometrischen Prinzipien beruhen. Der bevorzugte elektromechanische Detektor **18** ist der coulometrische Detektor, der aus einem Durchfluß- oder porösen Carbonsgraphit amperometrischen Detektor besteht, wobei die Probe durch die Elektrode hindurch passiert, was zu einer 100%-igen Nachweiseffizienz führt. Um jede Komponente vollständig nachzuweisen, wird ein Array von **16** coulometrischen Detektoren, die jeder bei einem unterschiedlichen Potential (im allgemeinen bei 60 mV anstiegen) gehalten wird. Beispiele von Herstellern dieses Typs von Detektor sind ESA (Bedford, MA) und Bioanalytical Systems Inc. (800-845-4246).

[0432] In einer alternativen Ausführungsform, die schematisch in **Fig. 16** dargestellt ist, sind die Probeneinführungsvorrichtung **12**, die Abtrennvorrichtung **14** und die Spaltvorrichtung **16** in Reihe wie oben diskutiert miteinander verbunden, um den Fluß der Probe aufrechtzuerhalten. Die Spaltvorrichtung **16** ist mit einer Nachweisvorrichtung **18** verbunden, die ein Massenspektrometer **40** oder ähnliches ist, die die Markierungen basierend auf dem Unterschied im Molekulargewicht zwischen jedem der Marker, die dazu verwendet wurden, um jede Art von DNA zu markieren, die in dem PCR-Schritt erzeugt wurde, nachzuweisen. Der Beste auf Unterschieden in der Masse basierende Detektor ist das Massenspektrometer. Für diese Verwendung wird das Massenspektrometer **40** typischerweise eine atmosphärischer Druckionisations (API)-Schnittstelle mit entweder Elektrospray- oder chemischer Ionisation, einen Quadrupol-Massenanalyser und einen Massenbereich von mindestens 50 bis 2600 m/z aufweisen. Beispiele von Herstellern eines geeigneten Massenspektrometers sind: Hewlett Packard (Palo Alto, CA), HP 1100 LC/MSD, Hitachi Instruments (San Jose, CA), M-1200H LC/MS, Perkin Elmer Corporation, Applied Biosystems Division (Foster City, CA) API 100 LC/MS oder API 300 LC/MS/MS, Finnigan Corporation (San Jose, CA) LCQ, MAT 95 S, Bruker Analytical Systems, Inc. (Billerica, MA) APEX, BioAPEX und ESQUIRE und Micromass (U.K.).

[0433] Die Nachweisvorrichtung **18** ist elektronisch mit einem Datenprozessor und Analysen **20** verbunden, der Daten von der Nachweisvorrichtung empfängt. Der Datenprozessor und der Analysen r schließen ein Softwareprogramm ein, das den nachgewiesenen Marken identifiziert. Der Datenprozessor und Analysen **230** in alternativen Ausführungsformen ist operativ mit der Injektionsvorrichtung **12**, der Trennvorrichtung **14**, den Fraktionssammler **15** und/oder der Spaltvorrichtung **16** verbunden, um die verschiedenen Komponenten des genetischen Fingerprinting-Systems **10** zu kontrollieren.

[0434] Das Software-Paket kartiert die elektrochemische Signatur eines bestimmten Markers mit einem spezifischen Primer und einer Retentionszeit. Software erzeugte Nukleinsäureprofile werden dann verglichen (Länge zu Länge, Fragment zu Fragment) und die Ergebnisse werden dem Verwender berichtet. Die Software wird sowohl Ähnlichkeit und Unterschiede in den Nukleinsäurefragmentprofilen hervorheben. Die Software wird auch in der Lage sein, die Sammlung von spezifischen Nukleinsäurefragmenten durch den Fraktions-sammler **15** zu steuern.

[0435] Das Software-Paket kartiert die m/z-Signatur eines bestimmten Markers mit einem bestimmten Primer und einer Retentionszeit. Software erzeugte Nukleinsäureprofile werden dann verglichen (Länge zu Länge, Fragment zu Fragment) und die Ergebnisse dem Verwender berichtet. Die Software wird sowohl die Ähnlichkeiten als auch die Unterschiede in den Nukleinsäurefragmentprofilen hervorheben. Die Software wird auch in der Lage sein, die Sammlung von spezifischen Nukleinsäurefragmenten durch den Fraktionssammler zu steuern.

[0436] Das System **18** in Übereinstimmung mit der vorliegenden Erfindung wird durch operatives miteinander Verbinden der multiplen Komponenten des Systems zur Verfügung gestellt. Demzufolge können eine oder mehrere Systemkomponenten, wie zum Beispiel die Probeneinführungsvorrichtung **12** und die Nachweisvorrichtung **18**, die in einem Labor in Betrieb sind, mit den anderen Komponenten des Systems, kombiniert werden (z. B. der Abtrennvorrichtung **14**, der Spaltvorrichtung **16** und dem Datenprozessor und Analyser **20**, um das Labor mit dem DNA-Sequenzierungssystem **10** der vorliegenden Erfindung auszurüsten.

[0437] Eine weitere Ausführungsform der vorliegenden Erfindung stellt ein Differential-Display-System zur Verfügung, das im allgemeinen aus einer Probeneinführungsvorrichtung, eine Vorrichtung, um die markierten Proben von Interesse zu trennen, einer Aufteilverrichtung, um eine variable Menge der Probe zu einem Fraktionssammler umzuleiten, eine Einheit, um die Marker von den Proben von Interesse abzuspalten, eine Vorrichtung zum Nachweis des Markers und ein Softwareprogramm, um die gesammelten Daten zu analysieren und sie in einem Differential-Displaymodus wiederzugeben, besteht. Es wird dem Fachmann ersichtlich sein, der sich in Besitz der vorliegenden Offenbarung befindet, daß die allgemeine Beschreibung viele Varianten für jede der aufgelisteten Komponenten haben kann. Das Differential-Displaysystem einer beispielhaften Ausführungsform

der vorliegenden Erfindung besteht aus ähnlichen Komponenten, die in **Fig. 15** dargestellt sind, einschließlich der Probeneinführvorrichtung **12**, der Abtrennvorrichtung **14** zur Abtrennung der Proben durch High-Performance-Liquid-Chromatographie (HPLC), der Abteilverrichtung **13**, dem Fraktionssammler **15**, der Photospaltungsanordnung **16**, um die Marker von den Proben von Interesse abzuspalten, der Nachweisvorrichtung **18** zum Nachweis der Marker durch Elektrochemie und den Datenprozessor und Analyser **20**, mit einem Softwareprogramm. Jede Komponente wird unten genauer diskutiert.

[0438] In dem Differential-Displaysystem nimmt die Probeneinführvorrichtung **12** automatisch einen gemessenen Aliquot **22** des PCR-Produkts, der in dem Differential-Displayverfahren erzeugt wurde und führt sie über eine herkömmliche Leitung **24** zu der Abtrennvorrichtung **14** (im allgemeinen ein HPLC) zu. Die Probeneinführvorrichtung **12** der beispielhaften Ausführungsform besteht aus einem Temperatur kontrollierten Autosampler **26**, der Mikrotiterplatten aufnehmen kann. Der Autosampler **26** muß Temperatur-kontrolliert sein, um die Integrität der erzeugten Nukleinsäureproben aufrechtzuerhalten und in der Lage sein 25 µl oder weniger an Probe zu injizieren. Hersteller dieses Typs von Produkt sind zum Beispiel durch Gilson (Middleton, WI) repräsentiert.

[0439] Die Probeneinführvorrichtung ist operativ in Reihe mit der Abtrennvorrichtung durch die herkömmliche Leitung **24** verbunden. Die PCR-Produkte in den gemessenen Aliquots **22**, die in der Auftrennvorrichtung **14** aufgenommen werden, werden temporal durch High-Performance-Liquid-Chromatographie abgetrennt, um abgetrennte DNA-Fragmente zur Verfügung zu stellen. Die High-Performance-Liquid-Chromatographie kann eine isokratische, binäre oder quaternäre Pumpe(n) **27** aufweisen und kann von zahlreichen Herstellern erworben werden (z. B. Hewlett Packard (Palo Alto, CA) HP 1100 oder 1090 Serien, Analytical Technology Inc. (Madison, WI), Perkin Elmer, Waters, usw.). Die Auftrennvorrichtung **14** schließt eine analytische HPLC-Säule **28** ein, die zur Verwendung zur Abtrennung der Oligonukleotide geeignet ist. Die Säule **28** ist eine analytische HPLC, zum Beispiel, nicht poröses Polystyroldivinylbenzol (2,2 µl Partikelgröße) fester Träger, modifiziert, die innerhalb eines pH-Bereichs von 2 bis 12, drücken von bis zu 3000 psi und einem Temperaturbereich von 10 bis 70°C betrieben werden kann. Eine Temperatur-Kontrollvorrichtung (z. B. ein Säulenofen) (nicht gezeigt) kann dazu verwendet werden, um die Temperatur der Säule zu kontrollieren. Solche Temperatur-Kontrollvorrichtungen sind im Stand der Technik bekannt und können zum Beispiel von Rainin Instruments (Tochterunternehmen von Varian Instruments, Palo Alto, CA) erworben werden. Eine geeignete Säule **28** ist unter dem kommerziellen Namen DNasep® erhältlich und ist von Serasep (San Jose, CA) erhältlich. Andere geeignete analytische HPLC-Säulen sind von anderen Herstellern erhältlich (z. B. Hewlett Packard (Palo Alto, CA), Beckman Industries (Brea, CA), Waters Corp. (Milford, MA) und Supelco (Bellefonte, PA)).

[0440] In der dargestellten Ausführungsform ist der Fraktionssammler **15** eine einzelne Komponente des Differential-Displaysystems **10**, die an die anderen Komponenten des Systems gekoppelt ist. In einer alternativen Ausführungsform ist der Fraktionssammler **15** in die Probeneinführungsvorrichtung **12** integriert. Demzufolge leitet die Fluß-Aufteilungsleitung **32** die getrennten Proben von dem Probenteiler **13** zurück zu der Probeneinführvorrichtung **12**.

[0441] Die Auftrennvorrichtung **14** in der dargestellten Ausführungsform enthält den Probenteiler **13**, der mit dem Flußstrom der Probe verbunden ist. Der Probenteiler **13** ist so angepaßt, um auf eine herkömmliche Weise veränderliche Mengen von Probe zu dem Fraktionssammler **15** entweder für die weitere Analyse oder Lagerung abzutheilen. Der Fraktionssammler **15** muß in der Lage sein, kleine Volumen aufzunehmen, eine Temperaturkontrolle zu niedrigen Temperaturen aufzuweisen und einstellbare Probenzeiten aufzuweisen. Hersteller von In-line-Abteilern schließen Upchurch (Oak Harbor, WA) ein.

[0442] Ein Fluß der getrennten DNA-Fragmente fließt durch eine herkömmliche Leitung **30** von der Auftrennvorrichtung **14** stromabwärts des Probenteilers **113** zu der Spaltvorrichtung **16**. Jedes der PCR-Produkte ist mit einem einmaligen spaltbaren (z. B. photospaltbaren) Marker markiert. Der Flußstrom von getrennten DNA-Fragmenten passiert durch oder entlang der Spaltvorrichtung **16**, wo der Marker zum Nachweis durch elektrochemischen Nachweis entfernt wird. In der beispielhaften Ausführungsform ist die Spaltvorrichtung **16** eine photospaltende Einheit, so daß der Flußstrom der Probe gegenüber einer ausgewählten Lichtenergie ausgesetzt wird. In einer Ausführungsform tritt die Probe in die photospaltende Einheit **16** ein und wird gegenüber der ausgewählten Lichtquelle für eine ausgewählte Zeitdauer ausgesetzt. In einer alternativen Ausführungsform wird der Strom an Probe in einem geeigneten Rohrteil oder ähnlichen in Angrenzung an die Lichtquelle entlang eines Weges geführt, der ein ausreichendes Aussetzen gegenüber der Lichtquelle zur Verfügung stellt, um die Marker von den abgetrennten DNA-Fragmenten abzuspalten. Eine photospaltende Einheit ist erhältlich von Supelco (Bellefonte, PA). Die Photospaltung kann bei vielen Wellenlängen mit einer Quecksilber/Xenon-Bogenlampe durchgeführt werden. Die Wellenlängengenauigkeit ist ungefähr 2 nm mit einer Bandbreite von 10 nm. Der bestrahlte Bereich ist kreisförmig und typischerweise ein Bereich von 10–100 Quadratzentimetern. In alternativen Ausführungsformen können andere Spaltvorrichtungen, die durch Säure, Base, Oxidation, Reduktion, Fluorid, Thiol austausch, Photolyse oder enzymatische Bedingungen spalten, dazu verwendet werden, die Marker von den abgetrennten DNA-Fragmenten zu entfernen.

[0443] Nachdem die Spaltvorrichtung **16** die Marker von den abgetrennten DNA-Fragmenten abspaltet, fließen die Marker durch ein konventionelles Röhrchen **32** zur der Nachweisvorrichtung **18** zum Nachweis jedes

Markers. Der Nachweis der Marker basiert auf dem Unterschied in dem elektrochemischen Potential zwischen jedem der Marker, die dazu verwendet werden, jede in dem PCR-Schritt erzeugte Art von DNA zu markieren. Der elektrochemische Detektor **18** kann auf entweder coulometrischen oder amperometrischen Prinzipien operieren. Der bevorzugte elektromechanische Detektor **18** ist der coulometrische Detektor, der aus einem Durchfluß oder porösen Carbonsgraphit amperometrischen Detektor besteht, wobei das Säuleneluat durch die Elektrode hindurch passiert, was zu einer 100%-igen Nachweiseffizienz führt. Um jede Komponente vollständig nachzuweisen, wird ein Array von 16 coulometrischen Detektoren, wobei jeder bei einem unterschiedlichen Potential gehalten wird (im allgemeinen bei 60 mV anstiegen) verwendet. Die Hersteller dieses Typs von Detektor schließen ESA (Bedford, MA) und Bioanalytical Systems Inc. (800-845-4246) ein.

[0444] In einer alternativen Ausführungsform des schematisch in **Fig. 16** dargestellten Differential-Displaysystems sind die Probeneinführungsvorrichtung **12**, die Abtrennvorrichtung **14** und die Spaltvorrichtung **16** in Reihe verbunden, wie oben diskutiert, um den Fluß an Probe aufrechtzuerhalten. Die Abspaltvorrichtung **16** ist mit einer Nachweisvorrichtung **18** verbunden, die die Marker basieren auf den Unterschied im Molekulargewicht zwischen jedem der Marker, die dazu verwendet werden, jede Art von in PCR-Schritt erzeugte DNA zu markieren, basiert. Der Beste auf Unterschieden in der Masse basierende Detektor ist das Massenspektrometer **40**. Für diese Verwendung wird das Massenspektrometer typischerweise eine atmosphärische Druckionisations (API)-Schnittstelle mit entweder Elektrospray oder chemischer Ionisation, einen Quadrupol-Massensanalyser und einen Massenbereich von mindestens 50 bis 2600 m/z aufweisen. Beispiele von Herstellern von geeigneten Massenspektrometern sind: Hewlett Packard (Palo Alto, CA), HP 1100 LC/MSD, Hitachi Instruments (San Jose, CA), M-1200H LC/MS, JEOL USA, Inc. (Peabody, MA), Perkin Elmer Corporation, Applied Biosystems Division (Foster City, CA) API 100 LC/MS oder API 300 LC/MS/MS, Finnigan Corporation (San Jose, CA) LCQ, MAT 95 S, MAT 95 Q, MAT 900 S, MAT 900 SQ und SSQ 7000, Bruker Analytical Systems, Inc. (Billerica, MA) APEX, BioAPEX und ESQUIRE.

[0445] Die Nachweisvorrichtung **18** ist elektronisch mit einem Datenprozessor und Analyser **20** verbunden, der Daten von der Nachweisvorrichtung empfängt. Der Datenprozessor und der Analyser **20** schließen ein Softwareprogramm ein, das den nachgewiesenen Marker und seine Position in der DNA-Sequenz identifiziert. Der Datenprozessor und Analyser **20** ist in alternativen Ausführungsformen operativ mit der Injektionsvorrichtung **12**, der Abtrennvorrichtung **14**, den Fraktionssammler **15** und/oder der Spaltvorrichtung **16** verbunden, um die verschiedenen Komponenten des Differential-Displaysystems zu kontrollieren.

[0446] Das Softwarepaket kartiert die Signatur eines gegebenen Markers auf einen spezifischen Primer und eine Retentionszeit. Software erzeugte Nukleinsäureprofile werden dann verglichen (Länge zu Länge, Fragment zu Fragment) und die Ergebnisse werden dem Verwender berichtet. Die Software wird sowohl Ähnlichkeiten als auch Unterschiede in den Nukleinsäurefragmentprofilen hervorheben. Die Software wird auch in der Lage sein, die Sammlung von spezifischen Nukleinsäurefragmenten durch den Fraktionssammler **15** zu steuern.

[0447] Das Softwarepaket kartiert die m/z-Signatur eines bestimmten Markers auf einen spezifischen Primer und einer Retentionszeit. Die Software erzeugten Nukleinsäureprofile werden dann verglichen (Länge zu Länge, Fragment zu Fragment) und die Ergebnisse dem Verwender berichtet. Die Software hebt sowohl Ähnlichkeiten als auch Unterschiede in den Nukleinsäurefragmentprofilen hervor. Die Software ist auch in der Lage, die Sammlung von spezifischen Nukleinsäurefragmenten durch den Fraktionssammler zu steuern.

[0448] Das Differential-Displaysystem wird durch operatives miteinander Verbinden der zahlreichen Komponenten des Systems zur Verfügung gestellt. Demzufolge können eine oder mehrere Komponenten des Systems, wie zum Beispiel die Probeneinführungsvorrichtung **12** und die Nachweisvorrichtung **18**, die in einem Labor in Betrieb sind, mit den anderen Komponenten des Systems kombiniert werden (z. B. der Abtrennvorrichtung **14**, der Spaltvorrichtung **16** und dem Datenprozessor und Analyser **20**, um das Labor mit einem System in Übereinstimmung mit der vorliegenden Erfindung auszurüsten.

[0449] Einzel-Nukleotid-Verlängerungsassay, Oligo-Ligationsassay oder Oligonukleotidsonde basierte Assaysysteme der vorliegenden Erfindung bestehen aus, im allgemeinen, einer Probeneinführungsvorrichtung, einer Vorrichtung um die markierten Proben von Interesse aufzutrennen, einer Vorrichtung um die Marker von den Proben von Interesse abzuspalten, einer Vorrichtung zum Nachweis des Markers und ein Softwareprogramm, um die gesammelten Daten zu analysieren. Es wird dem Fachmann ersichtlich sein, wenn in Besitz der vorliegenden Beschreibung, daß die allgemeine Beschreibung viele Varianten für jede der aufgelisteten Komponenten aufweisen kann. Wie am besten in **Fig. 17** gesehen wird (brauche Figur #) besteht ein bevorzugtes Einzel-Nukleotid-Verlängerungsarray, Oligoligationsassay oder Oligonukleotidsonde basiertes Assaysystem **200** aus einer Probeneinführungsvorrichtung **212**, einer Abtrennvorrichtung **214**, daß die Proben durch High-Performance-Liquid-Chromatographie (HPLC) trennt, einer Spaltvorrichtung **216**, um die Marker von den Proben von Interesse abzuspalten, einer Nachweisvorrichtung **218** zum Nachweis der Marker durch Massenspektrometrie und einen Datenprozessor und Analyser **220**, der ein Softwareprogramm enthält. Jede Komponente wird unten genauer diskutiert.

[0450] Die Probeneinführungsvorrichtung **212** nimmt automatisch ein gemessenes Aliquot **222** des durch eine Viel-

zahl von Verfahren (PCR, Ligationen, Verdau, Nuklearen, usw.) erzeugten Nukleinsäurefragments und führt sie über eine herkömmliche Leitung **224** zu einer Abtrennvorrichtung **214** (im allgemeinen ein HPLC) zu. Die Probenzuführvorrichtung **212** der beispielhaften Ausführungsform besteht aus einem Temperatur kontrollierten Autosampler **226**, der Mikrotiterplatten aufnehmen kann. Der Autosampler **226** muß Temperatur kontrolliert sein, um die Integrität der erzeugten Nukleinsäureproben aufrechtzuerhalten und in der Lage sein, 25 µl oder weniger an Probe zu injizieren. Hersteller dieses Produkts sind zum Beispiel durch Gilson (Middleton, WI) repräsentiert.

[0451] Die Probenzuführvorrichtung ist operativ in Reihe mit der Abtrennvorrichtung durch die herkömmliche Leitung **224** verbunden. Die Nukleinsäureprodukte (die durch PCR, Ligationsreaktionen, Verdau, Nukleasen usw. hergestellt werden können) in den gemessenen Aliquots **222**, die in der Auftrennvorrichtung **214** anfangen werden, werden temporal durch High-Performance-Liquid-Chromatographie aufgetrennt. Die High-Performance-Liquid-Chromatographie kann eine isokratische, binäre oder quaternäre Pumpe(n) **227** aufweisen und kann von zahlreichen Herstellern erworben werden (z. B. Hewlett Packard (Palo Alto, CA) HP 1100 oder 1090 Serie, Beckman Instruments Inc. (800-742-2435), Bioanalytical Systems, Inc. (800-845-4246), ESA, Inc. (508) 250-700), Perkin Elmer Corp. (800-762-4000), Varian Instruments (800-926-3000), Waters Corp. (800-254-4752)).

[0452] Die Abtrennvorrichtung **214** schließt eine analytische HPLC-Säule **228** ein, die zur Verwendung zur Abtrennung der Nukleinsäurefragmente geeignet ist. Die Säule **228** ist eine analytische HPLC, zum Beispiel, nicht poröses Polystyroldivinylbenzol (2,2 µm Partikelgröße) fester Träger, der innerhalb eines pH-Bereichs von 2 bis 12, drücken von bis zu 3000 psi und einem Temperaturbereich von 10 bis 70°C betrieben werden kann. Eine Temperatur-Kontrollvorrichtung (z. B. ein Säulenofen) (nicht gezeigt) kann dazu verwendet werden, die Temperatur der Säule zu kontrollieren. Solche Temperatur-Kontrollvorrichtungen sind im Stand der Technik bekannt und können zum Beispiel von Rainin Instruments (Tochterunternehmen von Varian Instruments, Palo Alto, CA) erworben werden. Eine geeignete Säule **228** ist unter dem kommerziellen Namen DNAsep® erhältlich und ist von Serasep (San Jose, CA) erhältlich. Eine große Vielzahl von HPLC-Säulen **228** kann für diese bestimmte technologische Einheit verwendet werden, da eine Einzel-Basenpaarauflösung nicht nötiger Weise erforderlich ist. Andere geeignete analytische HPLC-Säulen sind von anderen Herstellern (z. B. Hewlett Packard (Palo Alto, CA), Beckman Instruments, Inc. (Brea, CA) und Waters Corp. (Milford, MA)) erhältlich.

[0453] Ein Strom der getrennten DNA-Fragmente (z. B. Sequenzierungs-Reaktionsprodukte) fließt durch eine herkömmliche Leitung **230** von der Trennvorrichtung **214** zu der Spaltvorrichtung **216**. Jedes der DNA-Fragmente ist mit einem einmaligen spaltbaren (z. B. Photo-spaltbaren) Marker markiert. Der Flußstrom von getrennten DNA-Fragmenten passiert durch oder entlang der Spaltvorrichtung **216**, wobei der Marker zum Nachweis durch Massenspektrometrie oder mit einem elektrochemischen Detektor entfernt wird. Die Photospaltungseinheit ist von Supelco (Bellefonte, PA) erhältlich. Die Photospaltung kann bei vielfachen Wellenlängen mit einer Quecksilber/Xenon-Bogenlampe durchgeführt werden. Die Wellenlängengenauigkeit ist ungefähr 2 nm mit einer Bandbreite von 10 nm. Der bestrahlte Bereich ist kreisförmig und typischerweise ein Bereich von 10–100 Quadratzentimetern. In alternativen Ausführungsformen können andere Spaltvorrichtungen, die durch Säure, Base, Oxidation, Reduktion, Fluorid, Thiolaustausch, Photolyse oder enzymatische Bedingungen spalten, dazu verwendet werden, die Marker von den getrennten DNA-Fragmenten zu entfernen.

[0454] Nachdem die Spaltvorrichtung **216** die Marker von den getrennten DNA-Fragmenten absplattet, fließen die Marker durch eine herkömmliche Leitung **232** zur Nachweisvorrichtung **218** zum Nachweis jedes Markers. Der Nachweis der Marker kann auf dem Unterschied im Molekulargewicht zwischen jedem der Marker, die dazu verwendet wurden, jede Art von den in den verschiedenen Assay-Schritten erzeugten DNA zu markieren, basiert sein. Der beste Detektor, der auf Unterschieden in der Masse basiert, ist das Massenspektrometer. Für diese Verwendung wird das Massenspektrometer typischerweise ein atmosphärische Druck-Ionisations (API)-Schnittstelle mit entweder Elektrospray oder chemischer Ionisation, ein Quadrapol-Massenanalyser und einen Massenbereich von mindestens 50 bis 2600 m/z aufweisen. Beispiele von Herstellern eines geeigneten Massenspektrometers sind: Hewlett Packard (Palo Alto, CA), HP 1100 LC/MSD, Hitachi Instruments (San Jose, CA) M-1200H LC/MS, Perkin Elmer Corporation, Applied Biosystems Division (Foster City, CA) API 100 LC/MS oder API 300 LC/MS/MS, Finnigan Corporation (San Jose, CA) LCQ, Bruker Analytical Systems, Inc. (Billerica, MA), ESQUIRE und Micromers (U.K.).

[0455] In einer schematisch in **Fig. 16** dargestellten alternativen Ausführungsform sind die Probeneinführungsvorrichtung **212**, die Abtrennvorrichtung **214** und die Spaltvorrichtung **216** in Reihe wie oben diskutiert miteinander verbunden, um den Fluß der Probe aufrechtzuerhalten. Die Spaltvorrichtung **216** ist mit einer Nachweisvorrichtung **218** verbunden, die ein elektrochemischer Detektor **240** ist, der die Marker basierend auf dem Unterschied im elektrochemischen Potential zwischen jedem der Marker, der dazu verwendet wurde, jede Art von DNA, die in dem Sequenzierungs-Reaktionsschritt erzeugt wurde, zu markieren, nachweist. Der elektrochemische Detektor **240** der beispielhaften Ausführungsform kann auf entweder coulometrischen oder amperometrischen Prinzipien arbeiten. Der bevorzugte elektrochemische Detektor **240** ist der coulometrische Detektor, der aus einem Durchfluß- oder porösen Carbonsgraphit amperometrischen Detektor besteht, wobei das

Säulen-Eluat durch die Elektrode hindurch passiert, was zu einer 100% Nachweiseffizienz führt. Um jede Komponente voll nachzuweisen, wird ein Array von 16 coulometrischen Detektoren, die jeder bei einem unterschiedlichen Potential gehalten werden (im allgemeinen bei 60 mV anstiegen) verwendet. Beispiele von Herstellern dieses Typs von Detektoren sind ESA (Bedford, MA) und Bioanalytical Systems Inc. (800-845-4246). Zusätzliche Hersteller von elektrochemischen Detektoren können in der Liste von anderen Herstellern gefunden werden, die unten gefunden wird.

[0456] Der elektrochemische Detektor **240** ist elektrisch mit dem Datenprozessor und Analyser **220** mit dem oben diskutierten Softwarepaket verbunden. Das Softwarepaket kartiert die nachgewiesene Eigenschaft (z. B. die Masse oder elektrochemische Signatur) eines gegebenen Markers auf einem spezifischen Proben-ID. Die Software wird in der Lage sein, das Nukleinsäurefragment von Interesse zu identifizieren und die ID-Information in jeweilige Datenbanken zu laden.

[0457] Die hier beschriebenen Analysesysteme weisen zahlreiche Vorteile gegenüber den herkömmlichen Gel-basierten Systemen auf. Einer der prinzipiellen Vorteile ist, daß die Systeme vollautomatisiert werden können. Bei Verwendung eines HPLC-basierten Auftrennungssystems können die Proben automatisch in die HPLC initiiert werden, wohingegen Gel-basierte Systeme eine manuelle Beladung erfordern. Es gibt auch signifikante Zeitersparnisse, die bei der Aufbauzeit gefunden werden (keine Gelformen sind zu reinigen, keine Gele zu gießen) und der Analysezeit (mehr als 4 Stunden für ein großes Gel gegenüber viel kürzeren Zeiten (5 Minuten bis 1 Stunde) für eine HPLC-Analyse. Zusätzlich gibt es einen Probendurchsatzvorteil. Bei Verwendung der in dieser Erfindung beschriebenen Marker können viele Proben in einer Charge analysiert werden (potentiell 384 Proben/Spur), wohingegen die Gel-basierten Analysen oft auf die 4 erhältlichen Fluorophore oder eine Probe/Spur begrenzt sind. Die verwendeten Gele sind inhärent empfindlich und können leicht brechen oder eine Luftblase oder einen anderen Mangel enthalten, der das gesamte Gel oder einige Spuren unbrauchbar macht. HPLC-Säulen sind solide und, wenn prä-gepackt erworben, frei von Packdefekten, was einen konsistenten, im allgemeinen gleichmäßigen Auftrennungsweg erzeugt. Die HPLC-Systeme verleihen auch eine bessere Qualitätssicherung, in das interne Standards verwendet werden können, aufgrund der Reproduzierbarkeit der HPLC-Säulen. Die Gelqualität ist sowohl zwischen Gelen als auch innerhalb eines Gels konsistent, was die Verwendung von Standards nahe zu unmöglich macht. Zuletzt sind sowohl die Massenspektrometrie und elektrochemischen Detektoren sensitiver als die Detektoren, die in den Gel-basierten Systemen verwendet werden, was geringer Nachweisgrenzen und die Analyse von weniger Probe ermöglicht, was für nicht-PCR-basierte Analysen brauchbar wäre.

#### Markierte Proben in Array-basierten Assays

[0458] Arrays mit kovalent angebondenen Oligonukleotiden wurden hergestellt, die dazu verwendet wurden, um eine DNA-Sequenz-Analyse durch Hybridisierung durchzuführen (Southern et al., Genomics 13: 1008, 1992; Drmanac et al., Science 260: 1649, 1993), Expressionsprofile zu bestimmen, auch Mutationen zu screenen und ähnliches. Im allgemeinen verwendet der Nachweis für diese Tests fluoreszente oder radioaktive Markierungen. Fluoreszente Markierungen können am direktesten durch ihre Absorption und Fluoreszenz-Emissions-Wellenlängen und -intensitäten identifiziert und quantifiziert werden. Ein Mikroskop/Kameraaufbau, unter der Verwendung einer Fluoreszenzlichtquelle ist ein bequemes Mittel zum Nachweis von Fluoreszenzmarker. Radioaktive Marker können durch Standard-Autoradiographie, Phosphor-Image-Analyse oder CCD-Detektor visualisiert werden. Für solche Marker ist die Zahl von verschiedenen Reaktionen, die zu einer einzelnen Zeit nachgewiesen werden können, begrenzt. Zum Beispiel begrenzt die Verwendung von 4 fluoreszenten Molekülen, wie zum Beispiel herkömmlich bei der DNA-Sequenz-Analyse verwendet, die Analyse auf zwei Proben zur gleichen Zeit. Im wesentlichen, aufgrund dieser Begrenzung muß jede Reaktion einzeln untersucht werden, wenn diese Detektorverfahren verwendet werden.

[0459] Ein vorteilhafteres Verfahren zum Nachweis ermöglicht, daß Vereinigen der Probenreaktionen auf mindestens einem Array und den gleichzeitigen Nachweis der Produkte. Durch Verwendung eines Markers, sowie denjenigen wie hier beschrieben, der ein unterschiedliches Molekulargewicht oder andere physikalische Eigenschaft in jeder Reaktion aufweist, kann das gesamte Set von Reaktionsprodukten zusammen geerntet und analysiert werden.

[0460] Wie oben angegeben sind die oben beschriebenen Verfahren für eine Vielzahl von Zwecken anwendbar. Zum Beispiel können die Arrays von Oligonukleotiden dazu verwendet werden, um auf die Qualität von Herstellung von Arrays, zur Quantifizierung oder qualitativen Analyse von Nukleinsäuremolekülen, zum Nachweis von Mutationen, zum Nachweis von Expressionsmolekülen, für toxikologische Tests und ähnliches verwendet werden.

#### 1. Sondenquantifizierung oder -typisierung

[0461] In dieser Ausführungsform werden die Oligonukleotide pro Element in einem Array immobilisiert, wo-

bei jedes Oligonukleotid in dem Element eine verschiedene oder verwandte Sequenz ist. Bevorzugterweise besitzt jedes Element ein bekanntes oder verwandtes Set von Sequenzen. Die Hybridisierung einer markierten Sonde an solchen einen Array ermöglicht die Charakterisierung einer Sonde und die Identifizierung und Quantifizierung der Sequenzen, die in einer Sondenpopulation enthalten sind.

[0462] Ein generalisiertes Assay-Format, das in den bestimmten Anwendungen, die unten diskutiert werden, verwendet werden kann, ist ein Sandwich-Assay-Format. In diesem Format werden eine Vielzahl von Oligonukleotiden von bekannter Sequenz auf einem festen Substrat immobilisiert. Das immobilisierte Oligonukleotid wird dazu verwendet, um eine Nukleinsäure (z. B. RNA, rRNA, ein PCR-Produkt, fragmentierte DNA) zu fangen und dann wird eine Signalsonde an einen verschiedenen Teil der gefangenen Zielnukleinsäure hybridisiert. [0463] Ein weiteres generalisiertes Assay-Format ist ein sekundäres Nachweissystem. In diesem Format werden die Arrays dazu verwendet, um markierte Nukleinsäuren zu identifizieren und zu quantifizieren und die in einem primären Bindungsassay verwendet wurden. Zum Beispiel, falls ein Assay zu einer markierten Nukleinsäure führt, kann die Identität dieser Nukleinsäure durch Hybridisierung an einen Array bestimmt werden. Diese Assay-Formate sind insbesondere brauchbar, wenn mit abspaltbaren Massenspektrometrie-Markern kombiniert.

## 2. Mutationsnachweis

[0464] Mutationen, die ein einzelnes Nukleotid einschließen, können in einer Probe durch Scanning-Techniken identifiziert werden, die dafür geeignet sind, um vorher unbekannte Mutationen zu identifizieren oder durch Techniken, die so aufgebaut sind, um bekannte Sequenzvarianten nachzuweisen, zu unterscheiden oder zu quantifizieren. Verschiedene Scanning-Techniken für den Mutationsnachweis wurden entwickelt, basierend auf der Beobachtung, daß Heteroduplexe von fehlgepaarten komplementären DNA-Strängen, die von Wildtyp- und Mutantensequenzen abgeleitet sind, ein abnormales Wanderungsverhalten aufweisen.

[0465] Die hier beschriebenen Verfahren können zum Mutationsscreening verwendet werden. Eine Strategie zum Nachweis einer Mutation in einem DNA-Strang ist durch Hybridisierung der Testsequenz an Zielsequenzen, die Wildtyp- oder Mutantensequenzen sind. Eine fehlgepaarte Sequenz weist einen destabilisierenden Effekt auf die Hybridisierung von kurzen Oligonukleotidsonden an eine Zielsequenz auf (siehe Wetmur, Crit. Rev. Biochem. Mol. Biol. 26: 227, 1991). Die Test-Nukleinsäurequelle kann genomische DNA, RNA, cDNA oder die Amplifikation von jeder dieser Nukleinsäuren sein. Bevorzugterweise wird die Amplifikation von Testsequenzen zuerst durchgeführt, gefolgt von Hybridisierung mit kurzen Oligonukleotidsonden, die auf einem Array immobilisiert sind. Ein Amplifikationsprodukt kann auch viele mögliche Sequenzvarianten durch Bestimmen seines Hybridisierungsmusters an einen Array von immobilisierten Oligonukleotidsonden gescannt werden.

[0466] Ein Marker, wie zum Beispiel hier beschrieben, wird im allgemeinen in das finale Amplifikationsprodukt inkorporiert durch Verwendung eines markierten Nukleotids oder durch Verwendung eines markierten Primers. Das Amplifikationsprodukt wird denaturiert und an den Array hybridisiert. Nicht-gebundenes Produkt wird abgewaschen und an den Array gebundener Marker wird durch eine der Verfahren hier nachgewiesen. Zum Beispiel, wenn abspaltbare Massenspektrometriemarkers verwendet werden, können viele Produkte gleichzeitig nachgewiesen werden.

## 3. Expressionsprofile/Differential Display

[0467] Säuger, wie zum Beispiel der Mensch, weisen etwa 100.000 verschiedene Gene in ihrem Genom auf, von denen nur ein kleiner Anteil, möglicherweise 15%, in jeder individuellen Zelle exprimiert werden. Differential-Display-Techniken erlauben die Identifizierung von Genen, die für individuelle Zelltypen spezifisch sind. Kurz gesagt werden im Differential Display die 3'-terminalen Teile von mRNAs amplifiziert und auf der Basis ihrer Größe identifiziert. Unter der Verwendung eines Primers, der so aufgebaut ist, um an die 5'-Grenze poly(A)-Schwanzes für die reverse Transkription zu binden, gefolgt von Amplifikation der cDNA unter der Verwendung von stromaufwärts liegenden willkürlichen Sequenzprimern, werden mRNA-Subpopulationen erhalten.

[0468] Wie hier beschrieben, wird ein Hochdurchsatzverfahren zur Messung der Expression von zahlreichen Genen, z. B. 1–2.000, zur Verfügung gestellt. Innerhalb einer Ausführungsform der Erfindung werden Verfahren zur Analyse des Muster zur Genexpression von einer ausgewählten biologischen Probe zur Verfügung gestellt, umfassend die Schritte von (a) Amplifizieren von cDNA aus einer biologischen Probe unter der Verwendung von einem oder mehreren markierten Primern, wobei der Marker korrelativ mit einer bestimmten Nukleinsäuresonde ist und nachweisbar durch nicht-fluoreszente Spektrometrie oder Potentiometrie ist, (b) Hybridisieren von amplifizierten Fragmenten an ein Array von Oligonukleotiden, wie hier beschrieben, (c) Abwaschen von nicht-hybridisierten Material und (d) Nachweisen des Markers durch nicht-fluoreszente Spektrometrie oder Potentiometrie, und daraus Bestimmen des Musters von Genexpression der biologischen Probe. Marker-basiertes Differential Display, insbesondere unter der Verwendung von abspaltbaren Massenspektrometriemarkern auf festen Substraten ermöglicht die Charakterisierung von unterschiedlich exprimierten Genen.



## 4. Einzelnukleotidverlängerungsassay

[0469] Die Primerverlängerungstechnik kann für den Nachweis von einzelnen Nukleotidaustauschen in einem Nukleinsäure-Templat verwendet werden (Sokolov, *Nucleic Acids Res.* 18: 3671, 1989). Die Technik ist im allgemeinen auf den Nachweis von jeder einzelnen Basenmutation anwendbar (Kuppuswamy et al., *Proc. Natl. Acad. Sci. USA* 88: 1143–1147, 1991). Kurz gesagt hybridisierte dieses Verfahren zuerst einen Primer an einer Sequenz angrenzend an einen bekannten einzelnen Nukleotidpolymorphismus. Die geprimte DNA wird dann Bedingungen unterzogen, in denen eine DNA-Polymerase ein markiertes dNTP, typischerweise ein ddNTP, hinzufügt, falls die nächste Base in dem Templat komplementär zu dem markierten Nukleotid in dem Reaktionsgemisch ist. In einer Modifikation wird cDNA zuerst auf eine Sequenz von Interesse amplifiziert, die einen einzelnen Basen-Unterschied zwischen zwei Allelen enthält. Jedes amplifizierte Produkt wird dann auf die Anwesenheit, Abwesenheit oder relativen Menge von jedem Allel durch Annealing eines Primers analysiert, der eine Base 5' zu dem Polymorphismus ist und Verlängern durch eine markierte Base (im allgemeinen durch ein Didesoxynukleotid). Nur wenn die korrekte Base in der Reaktion erhältlich ist, wird eine Base am 3'-Ende des Primers inkorporiert werden. Die Verlängerungsprodukte werden dann durch Hybridisierung an einen Array von Nukleotiden analysiert, so daß ein nicht-verlängertes Produkt nicht hybridisieren wird.

[0470] Kurz gesagt, wird in der vorliegenden Erfindung jedes Didesoxynukleotid mit einem einmaligen Marker markiert. Von den vier Reaktionsgemischen wird nur eine einen Didesoxy-Terminator an die Primer-Sequenz anfügen. Falls die Mutation vorhanden ist, wird sie durch den einmaligen Marker, auf dem Didesoxynukleotid nach der Hybridisierung an den Array nachgewiesen werden. Multiple Mutationen können gleichzeitig durch Agieren des DNA-Primers mit einem einmaligen Marker auch bestimmt werden. Daher werden die DNA-Fragmente in vier verschiedenen Reaktion reagiert, die jede einen verschieden markierten Didesoxy-Terminator enthalten, wobei der Marker mit einem bestimmten Didesoxynukleotid korrelativ ist und die nicht-fluoreszente Spektrometrie oder Potentiometrie nachweisbar ist. Die DNA-Fragmente werden dann an einen Array hybridisiert und nicht-hybridisiertes Material wird weg gewaschen. Die Marker werden von den hybridisierten Fragmenten abgespalten und durch die jeweilige Nachweisttechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen. Die nachgewiesenen Marker können mit den zu untersuchenden bestimmten DNA-Fragment korreliert werden sowie der Identität des Mutanten-Nukleotids.

## 5. Oligonukleotid-Ligationsassay

[0471] Der Oligonukleotid-Ligationsassay, (OLA) (Landegren et al., *Science* 241: 487, 1988) wird für die Identifizierung von bekannten Sequenzen in sehr großen und komplexen Genomen verwendet. Das Prinzip der OLA-Reaktion beruht auf der Fähigkeit von Ligase, zwei diagnostische Oligonukleotide kovalent zu verknüpfen, wenn sie benachbart zueinander auf einem gegebenen DNA-Target hybridisieren. Wenn die Sequenzen an der Sondenverbindungsstelle nicht perfekt basengepaart sind, werden die Sonden nicht von der Ligase verknüpft werden. Bei der Anwendung von Markierungen sind diese an die Sonde angebracht, die an das amplifizierte Produkt ligiert wird. Nach Abschluß der OLA werden die Fragmente an einen Array von komplementären Sequenzen hybridisiert, die Markierungen abgespalten und mit Massenspektrometrie nachgewiesen.

[0472] Innerhalb einer Ausführungsform der Erfindung werden Verfahren zum Nachweis der Identität eines Nukleinsäuremoleküls oder zum Nachweis eines selektierenden Nukleinsäuremoleküls in, zum Beispiel, einer biologischen Probe, zur Verfügung gestellt, wobei die Technik des Oligonukleotid-Ligationsassays verwendet wird. Kurz gesagt umfassen solche Verfahren allgemein die Schritte von Durchführen einer Amplifikation auf der Ziel-DNAS, gefolgt von der Hybridisierung mit der 5'-Ende markierten Reporter-Sonde und einer 5'-phosphorylierten Sonde. Die Probe wird mit T4-DNA-Ligase inkubiert. Die DNA-Stränge mit ligierten Proben werden auf dem Array durch Hybridisierung an ein Array gefangen, wobei nicht-ligierte Produkte nicht hybridisieren. Die Marker werden von den abgetrennten Fragmenten abgespalten und dann werden die Marker durch die jeweilige Nachweisttechnologie (z. B. Massenspektrometrie, Infrarotspektrometrie, potentiostatische Amperometrie oder UV/sichtbare Spektrophotometrie) nachgewiesen.

## 6. Andere Assays

[0473] Die hier beschriebenen Verfahren können auch dazu verwendet werden, um zu genotypisieren oder zur Identifizierung von Viren oder Mikroben. Zum Beispiel können F + RNA-Coliphagen brauchbare Kandidaten als Indikatoren für enterische Virus-Kontamination sein. Die Genotypisierung durch Nukleinsäureamplifikations- und Hybridisierungsverfahren sind verlässliche, schnelle, einfache und billige Alternativen zur Serotypisierung (Kafatos et al., *Nucleic Acids Res.* 7: 1541, 1979). Die Amplifikationstechniken und Nukleinsäurehybridisierungstechniken wurden erfolgreich dazu verwendet, um eine Vielzahl von Mikroorganismen einschließlich *E. coli* (Feng, *Mol. Cell Probes* 7: 151, 1993), Rotavirus (Sethaburt et al., *J. Med Virol.* 37: 192, 1992), He-

patitis-C-Virus (Stuyver et al., J. Gen Virol. 74: 1093, 1993) und Herpes-Simplex-Virus (Matsumoto et al., J. Virol. Methods 40: 119, 1992) zu klassifizieren.

[0474] Genetische Änderungen wurden in einer Vielzahl von experimentellen Säuger- und menschlichen Neoplasmen beschrieben und stellen die morphologische Basis für die Sequenz von morphologischen Veränderungen dar, die bei der Karzinogenese beobachtet wurden (Vogelstein et al., NEJM 319: 525, 1988). In den letzten Jahren mit dem Auftauchen von molekularbiologischen Techniken, waren allelische Verluste auf bestimmten Chromosomen oder Mutationen von Tumorsuppressorgenen sowie Mutationen in bestimmten Onkogenen (z. B. c-myc, c-jun. und der ras-Familie) die am meisten untersuchten Gruppen. Vorherige Arbeit (Finkelstein et al., Arch Surg. 128: 526, 1993) identifizierte eine Korrelation zwischen verschiedenen spezifischen Typen von Punktmutationen in den K-ras-Onkogenen und der Phase bei der Diagnose in den colorectalen Karzinomen. Die Ergebnisse lassen vermuten, daß die Mutationsanalyse wichtige Informationen über die Tumorigenität zur Verfügung stellen könnte, einschließlich des Musters und der Verbreitung von Metastasen. Der prognostische Wert von TP53- und K-ras-2-Mutationsanalyse in Phase III-Karzinomen des Colons wurde kürzlich gezeigt (Pricolo et al., Am. J. Surg. 171: 41, 1996). Es ist daher ersichtlich, daß die Genotypisierung von Tumoren und prä-kanzerösen Zellen und der spezifische Mutationsnachweis bei der Behandlung von Krebsarten in Menschen zunehmend wichtig sein werden.

#### Markierte Sonden in Array-basierten Sonden

[0475] Die hier beschriebenen markierten Biomoleküle können dazu verwendet werden, um (nicht-markierte) Arrays von Biomolekülen zu untersuchen. Bevorzugte Arrays von Biomolekülen enthalten ein festes Substrat, das eine Oberfläche umfaßt, wobei die Oberfläche mindestens teilweise mit einer Lage von Poly(ethylenimin) (PEI) bedeckt ist. Die PEI-Lage umfaßt eine Vielzahl von diskreten ersten Regionen, die an eine kontinuierlich zweite Region angrenzen und umgeben sind. Die ersten Regionen werden durch die Anwesenheit eines Biomoleküls und PEI definiert, während die zweite Region durch die Anwesenheit von PEI und das wesentliche Fehlen des Biomoleküls definiert ist. Bevorzugterweise ist das Substrat eine Glasplatte oder ein Silizium-Wafer. Jedoch kann das Substrat zum Beispiel Quarz, Gold, Nylon-6,6, Nylon oder Polystyrol sowie Verbundmaterialien davon sein, wie oben beschrieben.

[0476] Die PEI-Beschichtung enthält bevorzugterweise PEI, das ein Molekulargewicht im Bereich von 100 bis 100.000 aufweist. Die PEI-Beschichtung kann direkt an das Substrat gebunden sein, unter der Verwendung von zum Beispiel silyliertem PEI. Alternativ kann ein Reaktionsprodukt eines bifunktionellen Kopplungsmittels zwischen der Substratoberfläche und der PEI-Beschichtung zwischengelagert sein, wobei das Reaktionsprodukt kovalent an sowohl die Oberfläche und die PEI-Beschichtung gebunden ist und die PEI-Beschichtung die Oberfläche sichert. Das bifunktionelle Kopplungsmittel enthält eine erste und eine zweite reaktive funktionelle Gruppe, wobei die erste reaktive funktionelle Gruppe zum Beispiel ein Tri(O-C<sub>1</sub>-C<sub>8</sub>alkyl)silan ist, und die zweite reaktive funktionelle Gruppe wie zum Beispiel ein Epoxid, Isocyanat, Isothiocyanat und Anhydridgruppe ist. Bevorzugte bifunktionelle Kopplungsmittel schließen 2-(3,4-Epoxycyclohexyl)ethyltrimethoxysilan; 3,4-Epoxybutyltrimethoxysilan; 3-Isocyanatopropyltriethoxysilan, 3-(triethoxysilyl)-2-methylpropylsuccinianhydrid und 3-(2,3-Epoxypropoxy)propyltrimethoxysilan ein.

[0477] Der Array der Erfindung enthält erste Biomolekül-enthaltende Regionen, wobei jede Region einen Bereich innerhalb des Bereichs von ungefähr 1.000 Quadratmikron bis ungefähr 100.000 Quadratmikron aufweist. In einer bevorzugten Ausführungsform weisen die ersten Regionen Bereiche auf, die von ungefähr 5.000 Quadratmikron bis ungefähr 25.000 Quadratmikron reichen.

[0478] Die ersten Regionen sind bevorzugterweise im wesentlichen kreisförmig, wobei die Kreise einen durchschnittlichen Durchmesser von ungefähr 10 Mikron bis 200 Mikron aufweisen. Ob kreisförmig oder nicht, sind die Grenzen der ersten Regionen im wesentlichen voneinander (durch die zweite Region) durch einen durchschnittlichen Abstand von mindestens ungefähr 25 Micron beabstandet, jedoch durch nicht mehr als ungefähr 1 cm (und bevorzugterweise durch nicht mehr als ungefähr 1.000 Mikron). In einem bevorzugten Array sind die Grenzen von benachbarten ersten Regionen durch eine durchschnittliche Distanz von ungefähr 25 Mikron bis ungefähr 100 Mikron beabstandet, wobei diese Distanz bevorzugterweise durch das Array hindurch konstant ist und die ersten Regionen sind bevorzugterweise in sich wiederholenden geometrischen Muster angeordnet, wie in den hier beigefügten Figuren gezeigt. In einem bevorzugten parallel wiederholenden geometrischen Muster sind alle benachbarten ersten Regionen durch ungefähr den selben Abstand (ungefähr 25 Mikron bis ungefähr 100 Mikron) getrennt.

[0479] In bevorzugten Ausführungsformen gibt es von 10 bis 50 erste Regionen auf dem Substrat. In anderen Ausführungsformen gibt es zwischen 50 und 400 erste Regionen auf dem Substrat. In noch einer anderen bevorzugten Ausführungsform gibt es 400–800 erste Regionen auf dem Substrat.

[0480] Das in den ersten Regionen lokalisierte Biomolekül ist bevorzugterweise ein Nukleinsäurepolymer. Ein bevorzugtes Nukleinsäurepolymer ist ein Oligonukleotid, das ungefähr 15 bis 50 Nukleotide aufweist. Das Biomolekül kann Amplifikationsreaktionsprodukte sein, die von ungefähr 50 bis 1.000 Nukleotide aufweisen. In

jeder ersten Region ist das Biomolekül bevorzugterweise bei einer durchschnittlichen Konzentration im Bereich von  $10^5$ – $10^9$  Biomoleküle pro 2.000 Quadratmikron einer ersten Region vorhanden. Weiter bevorzugt reicht die durchschnittliche Konzentration von Biomolekül von  $10^7$  bis  $10^9$  Biomolekülen pro 2.000 Quadratmikron. In der zweiten Region ist das Biomolekül bevorzugterweise bei einer durchschnittlichen Konzentration von weniger als  $10^3$  Biomolekülen pro 2.000 Quadratmikron der zweiten Region vorhanden und weiter bevorzugt bei einer durchschnittlichen Konzentration von weniger als  $10^2$  Biomolekülen pro 2.000 Quadratmikron. Am meisten bevorzugt enthalten die zweiten Regionen überhaupt kein Biomolekül.

[0481] Die zur Anheftung der Lage von PEI an das Substrat verwendete Chemie hängt zu einem wesentlichen Teil von der chemischen Identität des Substrats ab. Der Stand der Technik stellt zahlreiche Beispiele von geeigneten Chemien zur Verfügung, die PEI an einen festen Träger anheften können. Zum Beispiel, wenn das Substrat Nylon-6,6 ist, kann die PEI-Beschichtung durch die in Van Ness, J. et al. *Nucleic Acids Res.* 19: 3345–3350, 1991 und die PCT Internationale Publikation WO 94/00600, die beide durch Bezugnahme hierdurch aufgenommen sind, beschriebenen Verfahren aufgetragen werden. Wenn der feste Träger Glas oder Silizium ist, werden geeignete Verfahren zur Aufbringung einer Lage von PEI gefunden in, z. B. Wassermann, B.P. *Biotechnology and Bioengineering* XXII: 271–287, 1980; und D'Souza, S. F. *Biotechnology Letters* 8: 643–648, 1986.

[0482] Bevorzugterweise wird die PEI-Beschichtung kovalent an das feste Substrat angebracht. Wenn das feste Substrat Glas oder Silizium ist, kann die PEI-Beschichtung kovalent an das Substrat unter der Verwendung von silysierender Chemie angebunden werden. Zum Beispiel ist PEI, das reaktive Siloxy-Endgruppen aufweist, käuflich von Gelest, Inc. (Tullytown, PA) erhältlich. Solches reaktives PEI kann mit einem Glaträger oder Silizium-Wafer in Kontakt gebracht werden und nach sanftem Schütteln wird das PEI an das Substrat adherieren. Alternativ kann ein bifunktionelles silysierendes Reagenz verwendet werden. Gemäß diesem Verfahren wird das Glas- oder Silizium-Substrat mit dem bifunktionellen silysierenden Reagenz behandelt, um das Substrat mit einer reaktiven Oberfläche auszustatten. Das PEI wird dann mit der reaktiven Oberfläche in Kontakt gebracht und bindet kovalent an die Oberfläche durch das bifunktionelle Reagenz.

[0483] Die in das Array-Format platzierten Biomoleküle sind auch ursprünglich in einer sogenannten "Arraying-Lösung" vorhanden. Um Biomoleküle in diskreten Regionen auf dem PEI-beschichteten Substrat zu platzieren, enthält die Arraying-Lösung bevorzugterweise ein Verdickungsmittel bei einer Konzentration von ungefähr 35 Vol% bis 80 Vol%, basierend auf dem Gesamtvolumen der Zusammensetzung, ein Biomolekül, das bevorzugterweise ein Oligonukleotid ist, bei einer Konzentration, die von 0,001 µg/ml bis 10 µg/ml reicht und Wasser.

[0484] Die Konzentration des Verdickungsmittels ist 35% V/V bis 80% V/V für flüssige Verdickungsmittel, wie zum Beispiel Glycerin. Die bevorzugte Konzentration von Verdickungsmittel in der Zusammensetzung hängt zu einem Ausmaß von der Temperatur ab, bei der das Arraying durchgeführt wird. Je geringer die Arraying-Temperatur ist, desto geringer die Konzentration von Verdickungsmittel, das verwendet zu werden braucht. Die Kombination von Temperatur und flüssiger Verdickungsmittel-Kombination-Kontrolle ermöglicht es, Arrays auf den meisten Typen von festen Trägern (z. B. Glas, Wafern, Nylon 6/6, Nylonmembranen, usw.) herzustellen.

[0485] Die Anwesenheit eines Verdickungsmittels weist den zusätzlichen Vorteil von Ermöglichen der konkurrierenden Anwesenheit von niedrigen Konzentrationen von verschiedenen anderen Materialien auf, die in Kombination mit dem Biomolekül vorhanden sind. Zum Beispiel können 0,001% V/V bis 1% V/V von Detergenzien in der Arraying-Lösung vorhanden sein. Dies ist brauchbar, da PCR-Puffer eine kleine Menge von Tween-20 oder NP-40 enthält und es ist häufig wünschenswert, Probennukleinsäuren direkt aus einem PCR-Gefäß ohne vorherige Reinigung der Amplikons anzuordnen. Die Verwendung eines Verdickungsmittels ermöglicht die Anwesenheit von Salzen (zum Beispiel NaCl, KCl oder  $MgCl_2$ ), Puffern (zum Beispiel Tris) und/oder Chelat-bildenden Reagenzien (zum Beispiel EDTA) als auch in der Arraying-Lösung vorhanden zu sein. Die Verwendung eines Verdickungsmittels weist auch den weiteren Vorteil von Ermöglichen der Verwendung von kreuzvernetzenden Reagenzien und/oder organischen Lösungsmitteln auf, die in der Arraying-Lösung vorhanden sein können. Wie käuflich erhalten, werden kreuzvernetzende Reagenzien herkömmlich in organischen Lösungsmitteln wie zum Beispiel DMSO, DMF, NMP, Methanol, Ethanol und ähnliches gelöst. Herkömmlich verwendete organische Lösungen können bei Arraying-Lösungen der Erfindung in den Spiegeln von 0,05% bis 20% (V/V) verwendet werden, wenn Verdickungsmittel verwendet werden.

[0486] Im allgemeinen verleihen die Verdickungsmittel eine erhöhte Viskosität an die Arraying-Lösung. Wenn eine geeignete Viskosität in der Arraying-Lösung erhalten wird, hat der erste Tropfen im wesentlichen die selbe Größe, wie zum Beispiel der aufgebrauchte 100ste Tropfen. Wenn eine ungeeignete Viskosität in der Arraying-Lösung verwendet wird, sind die ersten aufgetragenen Tropfen signifikant größer als die späteren Tropfen, die aufgetragen werden. Die gewünschte Viskosität liegt zwischen derjenigen von reinem Wasser und reinem Glycerin.

[0487] Das Biomolekül in dem Array kann ein Nukleinsäurepolymer oder ein Analog davon sein, wie zum Beispiel PNA, Phosphorothioate und Methylphosphonate. Nukleinsäure betrifft sowohl Ribonukleinsäure als auch

Desoxyribonukleinsäure. Das Biomolekül kann unnatürliche und/oder synthetische Basen umfassen. Das Biomolekül kann ein Einzel- oder doppelsträngiges Nukleinsäurepolymer sein.

[0488] Ein bevorzugtes Biomolekül ist ein Nukleinsäurepolymer, das Oligonukleotide enthält (bis zu ungefähr 100 Nukleotidbasen) und Polynukleotide (über ungefähr 100 Basen). Ein bevorzugtes Nukleinsäurepolymer wird aus 15–50 Nukleotidbasen gebildet. Ein anderes bevorzugtes Nukleinsäurepolymer weist 50–1.000 Nukleotidbasen auf. Das Nukleinsäurepolymer kann ein PCR-Produkt, PCR-Primer oder Nukleinsäure-Duplex sein, um nur ein paar Beispiele aufzulisten. Jedoch kann im wesentlichen jeder Nukleinsäuretyp kovalent an eine PEI-beschichtete Oberfläche angebracht werden, wenn die Nukleinsäure ein primäres Amin enthält, wie unten diskutiert. Die typische Konzentration von Nukleinsäurepolymer in der Arraying-Lösung ist 0,001–10 µg/ml, bevorzugterweise 0,01–1 µg/ml und am meisten bevorzugt 0,05–0,5 µg/ml.

[0489] Bevorzugte Nukleinsäuremoleküle sind "Amin-modifiziert", in das sie modifiziert wurden um ein primäres Amin an ein 5'-Ende des Nukleinsäurepolymers zu enthalten, bevorzugterweise mit einer oder mehrerer Methyl- ( $-CH_2-$ ) Gruppen, angeordnet zwischen dem primären Amin und dem Nukleinsäureteil des Nukleinsäurepolymers. Sechs ist eine bevorzugte Zahl von Methylengruppen. Amin-modifizierte Nukleinsäurepolymere sind bevorzugt, da sie kovalent an einen festen Träger durch die 5'-Amingruppe gekoppelt werden können. PCR-Produkte können unter der Verwendung von 5'-Hexylamin-modifizierten PCR-Primern angeordnet werden. Die Nukleinsäure-Duplexe können nach der Einführung von Aminen durch Nick-Translation unter der Verwendung von Aminoallyl-dUTP (Sigma, St. Louis, MO) angeordnet werden. Die Amine können in Nukleinsäuren durch Polymerasen, wie zum Beispiel terminale Transferase und Aminoallyl-dUTP oder durch Ligation von kurzen Amin-enthaltenden Nukleinsäurepolymeren auf Nukleinsäuren durch Ligaren eingeführt werden.

[0490] Bevorzugterweise wird das Nukleinsäurepolymer aktiviert, bevor es mit der PEI-Beschichtung in Kontakt gebracht wird. Dies kann bequem durch Kombinieren von Aminfunktionalisiertem Nukleinsäurepolymer mit einer multi-funktionellen Amin-reaktiven Chemikalie, wie zum Beispiel Trichlortriazin erreicht werden. Wenn das Nukleinsäurepolymer eine 5'-Aminogruppe enthält, kann diese 5'-Amin mit Trichlortriazin reagiert werden, auch bekannt als Cyanurchlorid (Van Ness et al., *Nucleic Acids Res.* 19(2): 3345–3350, 1991). Bevorzugterweise wird ein Überschuß von Cyanurchlorid zu der Nukleinsäurepolymerlösung hinzugefügt, wobei ein 10- bis 1.000-facher molarer Überschuß von Cyanurchlorid über die Zahl von Aminen in dem Nukleinsäurepolymer in der Arraying-Lösung bevorzugt ist. Auf diese Weise hat die Hauptzahl von Amin-terminierten Nukleinsäurepolymeren mit einem Molekül von Trichlortriazin reagiert, so daß das Nukleinsäurepolymer mit Trichlortriazin beendet wird.

[0491] Bevorzugterweise wird die Arraying-Lösung unter der Verwendung eines herkömmlichen Puffers, wie zum Beispiel Natriumphosphat, Natriumborat, Natriumcarbonat und Tris HCl gepuffert. Ein bevorzugter pH-Bereich für die Arraying-Lösung ist 7 bis 9, wobei ein bevorzugter Puffer frisch präpariertes Natriumborat bei pH 8,3 bis pH 8,5 ist. Um eine typische Arraying-Lösung herzustellen, wird Hexylamin-modifiziertes Nukleinsäurepolymer in 0,2 M Natriumborat, pH 8,3, bei 0,1 µg/ml auf ein Gesamtvolumen von 50 µl plaziert. 10 µl einer 15 mg/ml Lösung von Cyanurchlorid werden dann hinzugefügt und der Reaktion wird ermöglicht, für eine Stunde bei 25°C mit konstantem Schütteln fortzuschreiten. Glycerin (Gibco Brl®, Gran Island, NY) wird bis zu einer finalen Konzentration von 56% hinzugefügt.

[0492] Die biomolekularen Arraying-Lösungen können auf die PEI-Beschichtung durch jede einer Zahl von Techniken, die im Augenblick bei der Mikrofabrikation verwendet werden, aufgetragen werden. Zum Beispiel können die Lösungen in einem Inkjet-Druckkopf plaziert werden und aus solch einem Kopf auf die Beschichtung ausgestoßen werden.

[0493] Ein bevorzugter Ansatz zur Zuführung von biomolekularer Lösung auf die PEI-Beschichtung verwendet eine modifizierte Springprobe. Springproben sind von verschiedenen Verkäufern erhältlich, einschließlich Everett Charles (Pomona, CA), Interconnect Devices Inc. (Kansas City, Kansas) und Test Connections Inc., (Upland, CA). Damit die wie oben beschriebenen erhältlichen Springsonden zufriedenstellend als flüssige Zubringvorrichtungen gemäß der vorliegenden Erfindung funktionieren, müssen ungefähr 1/1000 bis 5/1000 eines Inchs von Metallmaterial von der Spitze der Sonde entfernt werden. Das Verfahren muß zu einer flachen Oberfläche führen, die senkrecht zu der Längsachse der Springsonde ist. Das Entfernen von ungefähr 1/1000 bis 5/1000 eines Inchs von Material von dem Boden der Spitze ist bevorzugt und kann leicht mit einem sehr feingekörnten Schleifstein durchgeführt werden. Spezifische Springsonden, die käuflich erhältlich sind und die modifiziert werden können, um eine planare Spitze wie oben beschrieben zur Verfügung zu stellen, schließen die XP54-Sonde ein, die durch Ostby Barton (eine Abteilung von Everett Charles (Pomona, CA)); die SPA-Sonde 25P, die durch Everett Charles hergestellt wird (Pomona, CA) und 43-P-"fluted"-Springsonde von Test Connections Inc., (Upland, CA).

[0494] Die wie oben beschriebenen Arraying-Lösungen können direkt in einem Arraying-Verfahren verwendet werden. Gemeint ist, die aktivierten Nukleinsäurepolymere brauchen nicht vor dem Druckschritt von nicht reagiertem Cyanurchlorid weg gereinigt zu werden. Typischerweise wird der Reaktion, die die aktivierte Nukleinsäure an den festen Träger anbringt ermöglicht, für 1 bis 20 Stunden bei 20 bis 50°C fortzuschreiten. Bevorzugterweise ist die Reaktionszeit 1 Stunde bei 25°C.

[0495] Die hier beschriebenen Anays sind besonders brauchbar bei der Durchführung von Hybridisierungstests, zum Beispiel unter der Verwendung von CMST-markierten Sonden. Jedoch müssen die Amine auf dem festen Träger, um solche Tests durchzuführen, vor der Durchführung des Hybridisierungsschrittes geschützt werden. Dies kann durch Reagieren des festen Trägers mit 0,1–2,0 M Succinianhydrid erreicht werden. Die bevorzugten Reaktionsbedingungen sind 1,0 M Succinianhydrid in 70% m-Pyrol und 0,1 M Natriumborat. Der Reaktion wird typischerweise erlaubt für 15 Minuten bis 4 Stunden zu erfolgen, mit einer bevorzugten Reaktionszeit von 30 Minuten bei 25°C. Restliches Succinianhydrid wird mit einer 3x-Wasser-Waschung entfernt.

[0496] Der feste Träger wird dann mit einer Lösung inkubiert, die 0,1–5 M Glycin in 0,1–10,0 M Natriumborat bei pH 7–9 enthält. Dieser Schritt "schützt" jedes Dichlortriazin, das kovalent an die PEI-Oberfläche gebunden sein kann, durch Überführung in Monochlortriazin. Die bevorzugten Bedingungen sind 0,2 M Glycin in 0,1 M Natriumborat bei pH 8,3. Der feste Träger kann dann Detergenz-enthaltenen Lösungen gewaschen werden, um nicht-gebundene Materialien zu entfernen, zum Beispiel Spuren von NMP. Bevorzugterweise wird der feste Träger auf 95°C in 0,01 M NaCl, 0,05 EDTA und 0,1 M Tris, pH 8,0 für 5 Minuten erhitzt. Dieser Erhitzungsschritt entfernt nicht-kovalent angebrachte Nukleinsäurepolymere, wie zum Beispiel PCR-Produkte, in dem Fall wo Doppelstrang-Nukleinsäure angeordnet wird, hat dieser Schritt auch den Effekt von Überführen des Doppelstrangs in die einzelsträngige Form (Denaturierung).

[0497] Die Arrays können dann durch Sonden (z. B. Oligonukleotide, Nukleinsäurefragmente, PCR-Produkte, usw.) untersucht werden, die mit zum Beispiel CMST-Markern, wie hier beschrieben, Radioisotopen, Fluorophoren oder Biotin markiert sein können. Die Verfahren zur Biotinylierung von Nukleinsäuren sind im Stand der Technik gut bekannt und werden geeignet durch Pierce (Avidin-Biotin Chemistry: A Handbook, Pierce Chemical Company, 1992, Rockford Illinois) beschrieben. Die Sonden werden im allgemeinen bei 0,1 ng/ml bis 10 µg/ml in Standard-Hybridisierungslösungen verwendet, die GuSCN, GuHCl, Formamid, usw. (siehe Van Ness und Chen, Nucleic Acids Res. 19: 5143–5151, 1991) enthalten.

[0498] Um das Hybridisierungsereignis (d. h. die Anwesenheit des Biotins) nachzuweisen wird der feste Träger mit Streptavidin-/Rettichperoxidase-Konjugat inkubiert. Solche Enzymkonjugate sind käuflich erhältlich von, zum Beispiel, Vector Laboratories (Burlingham, CA). Das Streptavidin bindet mit hoher Affinität an das Biotinmolekül, was die Rettichperoxidase in die Nähe der hybridisierten Sonde bringt. Nicht-gebundenes Streptavidin-/Rettichperoxidase-Konjugat wird in einem einzelnen Waschschrift weg gewaschen. Die Anwesenheit von Rettichperoxidaseenzym wird dann unter der Verwendung eines ausfällenden Substrats in der Anwesenheit von Peroxid und den geeigneten Puffern nachgewiesen.

[0499] Ein blaues Enzymprodukt, das auf einer reflektierenden Oberfläche, wie zum Beispiel einem Wafer aufgetragen wird, weist einen vielfach niedrigeren Spiegel des Nachweises (LLD), verglichen mit demjenigen, der für ein colorimetrischen Substrat erwartet wird, auf. Weiterhin ist die LLD für verschieden gefärbte Enzymprodukte sehr verschieden. Zum Beispiel ist die LLD für 4-Methoxynaphthol (was ein ausgefälltes blaues Produkt produziert) pro 50 µM Durchmesser-Fleck ungefähr 1.000 Moleküle, wohingegen ein rot ausfallendes Substrat eine LLD von ungefähr 1.000-fach höher bei 1.000.000 Moleküle pro 50 µM Durchmesser-Fleck ergibt. Die LLD wird durch Untersuchen der Oberfläche mit einem Mikroskop (wie zum Beispiel des Axiotech-Mikroskop, käuflich erhältlich von Zeiss), ausgerüstet mit einer sichtbaren Lichtquelle und einer CCD-Kamera (Princeton Instruments, Princeton, NJ) nachgewiesen. Ein Bild von ungefähr 10.000 µM × 10.000 µM kann zur selben Zeit gescannt werden.

[0500] Um das blaue colorimetrische Nachweisschema zu verwenden, muß die Oberfläche nach der enzymatischen Reaktion wieder rein sein und der Wafer oder der Glasträger muß in einem trockenen Zustand gescannt werden. Zusätzlich muß die enzymatische Reaktion vor der Sättigung der Referenzflecken gestoppt werden. Für Rettichperoxidase ist dies ungefähr 2–5 Minuten.

[0501] Es ist ebenfalls möglich, chemilumineszente Substrate für alkalische Phosphatase oder Rettichperoxidase (HRP) oder Fluoreszenz-Substrate für HRP oder alkalische Phosphatase zu verwenden. Beispiele schließen die Dioxethan-Substrate für alkalische Phosphatase, erhältlich von Perkin Elmer oder Attophos HRP-Substrate von JBL Scientific (San Luis, Obispo, CA) ein.

[0502] Die folgenden Beispiele werden zur Veranschaulichung und nicht zur Beschränkung angeboten.

[0503] Sofern nicht anders angegeben, können Chemikalien, wie sie in den Beispielen verwendet werden, von Aldrich Chemical Company, Milwaukee, WI erhalten werden. Die folgenden Abkürzungen, mit den angegebenen Bedeutungen, werden hierin verwendet.

ANP = 3-(Fmoc-Amino)-3-(2-nitrophenyl)propionsäure

NBA = 4-(Fmoc-Aminomethyl)-3-nitrobenzoesäure

HATU = O-7-Azabenzotriazol-1-yl-N,N,N',N'-tetramethyluroniumhexafluorophosphat

DIEA = Diisopropylethylamin

MCT = Monochlortriazin

NMM = 4-Methylmorpholin

NMP = N-Methylpyrrolidon

ACT375 = ACT375-Peptidsynthesizer von Advanced ChemTech, Inc., Louisville, KY

ACT = Advanced ChemTech, Inc., Louisville, KY  
 NovaBiochem = CalBiochem-NovaBiochem International, San Diego, CA  
 TFA = Trifluoressigsäure  
 Tfa = Trifluoracetyl  
 iNIP = N-Methylisonipecotinsäure  
 Tfp = Tetrafluorphenyl  
 DIAEA = 2-(Diisopropylamino)ethylamin  
 MCT = Monochlortriazen  
 5'-AH-ODN = Oligodesoxynukleotid mit 5'-Aminohexyl-Schwanz

## BEISPIELE

## BEISPIEL 1

## HERSTELLUNG VON SÄURELABILEN LINKERN ZUR VERWENDUNG BEI DER SEQUENZIERUNG EINES SPALTBAREN MG-IDENTIFIKATORS

A. Synthese von Pentafluorphenylestern von chemisch spaltbaren Massenspektroskopiemarkierungen, um Markierungen mit Carbonsäureamid-Enden freizusetzen

[0504] **Fig. 1** zeigt das Reaktionsschema.

[0505] Schritt A. TentaGel-S-AC-Harz (Verbindung II; erhältlich von ACT; 1 Äq.) wird mit DMF im Sammelgefäß des ACT357-Peptidsynthesizers (ACT) suspendiert. Verbindung I (3 Äq.), HATU (3 Äq.) und DIEA (7,5 Äq.) in DMF werden zugegeben und das Sammelgefäß für 1 h geschüttelt. Das Lösungsmittel wird entfernt und das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Die Kopplung von I an das Harz und die Waschschritte werden wiederholt, um Verbindung III zu ergeben.

[0506] Schritt B. Das Harz (Verbindung III) wird mit 25%igem Piperidin in DMF vermischt und für 5 min geschüttelt. Das Harz wird filtriert, anschließend mit 25%igem Piperidin in DMF vermischt und für 10 min geschüttelt. Das Lösungsmittel wird entfernt, das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen und direkt in Schritt C verwendet.

[0507] Schritt C. Das Harz ohne Schutzgruppe aus Schritt B wird in DMF suspendiert und dazu werden eine Aminosäure mit Fmoc-Schutzgruppe, die Aminofunktionalität in ihrer Seitenkette enthält (Verbindung IV, z. B. Alpha-N-Fmoc-3-(3-Pyridyl)-alanin (erhältlich von Synthetech, Albany, OR; 3 Äq.), HATU (3 Äq.) und DIEA (7,5 Äq.) in DMF zugegeben. Das Gefäß wird für 1 h geschüttelt. Das Lösungsmittel wird entfernt und das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Die Kopplung von IV an das Harz und die Waschschritte werden wiederholt, um Verbindung V zu ergeben.

[0508] Schritt D. Das Harz (Verbindung V) wird mit Piperidin behandelt, wie in Schritt B beschrieben, um die Fmoc-Gruppe zu entfernen. Das Harz ohne Schutzgruppe wird anschließend gleichmäßig vom ACT357 aus dem Sammelgefäß in 16 Reaktionsgefäße aufgeteilt.

[0509] Schritt E. Die 16 Aliquote aus Harz ohne Schutzgruppe aus Schritt D werden in DMF suspendiert. Zu jedem Reaktionsgefäß werden die geeignete Carbonsäure  $VI_{1-16}$  ( $R_{1-16}CO_2H$ ; 3 Äq.), HATU (3 Äq.) und DIEA (7,5 Äq.) in DMF zugegeben. Die Gefäße werden für 1 h geschüttelt. Das Lösungsmittel wird entfernt und die Harz-Aliquote mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Die Kopplung von  $VI_{1-16}$  an die Harz-Aliquote und die Waschschritte werden wiederholt, um die Verbindungen  $VII_{1-16}$  zu ergeben.

[0510] Schritt F. Die Harz-Aliquote (Verbindungen  $VII_{1-16}$ ) werden mit  $CH_2Cl_2$  (3X) gewaschen. Zu jedem der Reaktionsgefäße wird 1% TFA in  $CH_2Cl_2$  zugegeben und die Gefäße für 30 min geschüttelt. Das Lösungsmittel wird aus den Reaktionsgefäßen in einzelne Röhren filtriert. Die Harz-Aliquote werden mit  $CH_2Cl_2$  (2X) und MeOH (2X) gewaschen und die Filtrate in die einzelnen Röhren hinein zusammengeführt. Die einzelnen Röhren werden im Vakuum eingedampft, was die Verbindungen  $VIII_{1-16}$  liefert.

[0511] Schritt G. Jede der freien Carbonsäuren  $VIII_{1-16}$  wird in DMF gelöst. Zu jeder Lösung wird Pyridin (1,05 Äq.) zugegeben, gefolgt von Pentafluorphenyltrifluoracetat (1,1 Äq.). Die Mischungen werden für 45 min bei Raumtemperatur gerührt. Die Lösungen werden mit EtOAc verdünnt, mit 1 M wäßriger Zitronensäure (3X) und 5% wäßrigem  $NaHCO_3$  (3X) gewaschen, über  $Na_2SO_4$  getrocknet, filtriert und in vacuo eingedampft, was die Verbindungen  $IX_{1-16}$  liefert.

B. Synthese von Pentafluorphenylestern von chemisch spaltbaren Massenspektroskopiemarkierungen, um Markierungen mit Carbonsäureenden freizusetzen

[0512] **Fig. 2** zeigt das Reaktionsschema.

[0513] Schritt A. 4-(Hydroxymethyl)phenoxybuttersäure (Verbindung I; 1 Äq.) wird zusammengebracht mit

DIEA (2,1 Äq.) und Allylbromid (2,1 Äq.) in  $\text{CHCl}_3$  und für 2 h unter Rückfluß erhitzt. Die Mischung wird mit EtOAc verdünnt, mit 1 N HCl (2X), pH 9,5 Carbonatpuffer (2X) und Salzlösung (1X) gewaschen, über  $\text{Na}_2\text{SO}_4$  getrocknet und in vacuo eingedampft, um den Allylester von Verbindung I zu ergeben.

[0514] Schritt B. Der Allylester von Verbindung I aus Schritt A (1,75 Äq.) wird in  $\text{CH}_2\text{Cl}_2$  mit einer Aminosäure mit Fmoc-Schutzgruppe, die Aminofunktionalität in ihrer Seitenkette enthält (Verbindung II, z. B. Alpha-N-Fmoc-3-(3-Pyridyl)-alanin, erhältlich von Synthetech, Albany, OR; 1 Äq.), N-Methylmorpholin (2,5 Äq.) und HATU (1,1 Äq.) zusammengebracht und bei Raumtemperatur für 4 h gerührt. Die Mischung wird mit  $\text{CH}_2\text{Cl}_2$  verdünnt, mit 1 M wäßriger Zitronensäure (2X), Wasser (1X) und 5%iger wäßriger  $\text{NaHCO}_3$  (2X) gewaschen, über  $\text{Na}_2\text{SO}_4$  getrocknet und in vacuo eingedampft. Verbindung III wird durch Flashchromatographie ( $\text{CH}_2\text{Cl}_2 \rightarrow \text{EtOAc}$ ) isoliert.

[0515] Schritt C. Verbindung III wird in  $\text{CH}_2\text{Cl}_2$  gelöst,  $\text{Pd}(\text{PPh}_3)_4$  (0,07 Äq.) und N-Methylalanin (2 Äq.) werden hinzugegeben und die Mischung wird bei Raumtemperatur für 4 h gerührt. Die Mischung wird mit  $\text{CH}_2\text{Cl}_2$  verdünnt, mit 1 M wäßriger Zitronensäure (2X) und Wasser (1X) gewaschen, über  $\text{Na}_2\text{SO}_4$  getrocknet und in vacuo eingedampft. Verbindung IV wird durch Flashchromatographie ( $\text{CH}_2\text{Cl}_2 \rightarrow \text{EtOAc} + \text{HOAc}$ ) isoliert.

[0516] Schritt D. TentaGel-S-AC-Harz (Verbindung V; 1 Äq.) wird mit DMF im Sammelgefäß des ACT357-Peptidsynthesizers (Advanced ChemTech Inc. (ACT), Louisville, KY) suspendiert. Verbindung IV (3 Äq.), HATU (3 Äq.) und DIEA (7,5 Äq.) in DMF werden hinzugefügt und das Sammelgefäß für 1 h geschüttelt. Das Lösungsmittel wird entfernt und das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Die Kopplung von IV an das Harz und die Waschschrte werden wiederholt, um Verbindung VI zu ergeben.

[0517] Schritt E. Das Harz (Verbindung VI) wird mit 25% Piperidin in DMF vermischt und für 5 min geschüttelt. Das Harz wird filtriert, dann mit 25% Piperidin in DMF vermischt und für 10 min geschüttelt. Das Lösungsmittel wird entfernt und das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Das Harz ohne Schutzgruppe wird dann gleichmäßig durch den ACT357 aus dem Sammelgefäß in 16 Reaktionsgefäße aufgeteilt.

[0518] Schritt F. Die 16 Aliquote aus Harz ohne Schutzgruppe aus Schritt E werden in DMF suspendiert. Zu jedem Reaktionsgefäß werden die geeignete Carbonsäure  $\text{VII}_{1-16}$  ( $\text{R}_{1-16}\text{CO}_2\text{H}$ ; 3 Äq.), HATU (3 Äq.) und DIEA (7,5 Äq.) in DMF zugegeben. Die Gefäße werden für 1 h geschüttelt. Das Lösungsmittel wird entfernt und die Harz-Aliquote mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Die Kopplung von  $\text{VII}_{1-16}$  an die Harz-Aliquote und die Waschschrte werden wiederholt, um die Verbindungen  $\text{VIII}_{1-16}$  zu ergeben.

[0519] Schritt G. Die Harz-Aliquote (Verbindungen  $\text{VIII}_{1-16}$ ) werden mit  $\text{CH}_2\text{Cl}_2$  (3X) gewaschen. Zu jedem der Reaktionsgefäße wird 1% TFA in  $\text{CH}_2\text{Cl}_2$  zugegeben und die Gefäße werden für 30 min geschüttelt. Das Lösungsmittel wird aus den Reaktionsgefäßen in einzelne Röhren filtriert. Die Harz-Aliquote werden mit  $\text{CH}_2\text{Cl}_2$  (2X) und MeOH (2X) gewaschen und die Filtrate in die einzelnen Röhren hinein zusammengegeben. Die einzelnen Röhren werden im Vakuum verdampft, was die Verbindungen  $\text{IX}_{1-16}$  liefert.

[0520] Schritt H. Jede der freien Carbonsäuren  $\text{IX}_{1-16}$  wird in DMF gelöst. Zu jeder Lösung wird Pyridin (1,05 Äq.) hinzugegeben, gefolgt von Pentafluorphenyltrifluoracetat (1,1 Äq.). Die Mischungen werden für 45 min bei Raumtemperatur gerührt. Die Lösungen werden mit EtOAc verdünnt, mit 1 M wäßriger Zitronensäure (3X) und 5% wäßriger  $\text{NaHCO}_3$  (3X) gewaschen, über  $\text{Na}_2\text{SO}_4$  getrocknet, filtriert und in vacuo eingedampft, was die Verbindungen  $\text{X}_{1-16}$  liefert.

## BEISPIEL 2

### NACHWEIS PHOTOLYTISCHER SPALTUNG VON T-L-X

[0521] Eine T-L-X-Verbindung, wie hergestellt in Beispiel 11, wurde mit Licht im nahen UV-Bereich für 7 min bei Raumtemperatur bestrahlt. Eine Rayonett-Fluoreszenz-UV-Lampe (Southern New England Ultraviolet Co., Middletown, CT) mit einem Emissionspeak bei 350 nm wird als eine UV-Lichtquelle verwendet. Die Lampe wird in einem Abstand von 15 cm von den Petrischalen mit Proben gestellt. SDS-Gelelektrophorese zeigt, daß >85% des Konjugats unter diesen Bedingungen gespalten werden.

## BEISPIEL 3

### HERSTELLUNG VON FLUORESCENZMARKIERTEN PRIMERN UND NACHWEIS DER SPALTUNG VON FLUOROPHOR (betrifft nicht den beanspruchten Gegenstand)

#### Synthese und Reinigung von Oligonukleotiden

[0522] Die Oligonukleotide (ODNs) werden auf automatisierten DNA-Synthesizern unter Verwendung der Standard-Phosphoramidit-Chemie, geliefert vom Anbieter, oder der H-Phosphonat-Chemie (Glen Research Sterling, VA) hergestellt. In geeigneter Weise blockierte dA-, dG-, dC- und T-Phosphoramidite sind in diesen Formen kommerziell erhältlich und synthetische Nukleoside können leicht in die geeignete Form überführt wer-

den. Die Oligonukleotide werden unter Verwendung der Standard-Phosphoramidit-Chemie, geliefert vom Anbieter, oder der H-Phosphonatchemie hergestellt. Oligonukleotide werden durch Adaptionen von Standardverfahren gereinigt. Oligonukleotide mit 5'-Tritylgruppen werden auf HPLC unter Verwendung einer 12 Mikrometer, 300 # Rainin (Emeryville, CA) Dynamax C-8 4,2 × 250 mm Umkehrphasensäule unter Verwendung eines Gradienten von 15% bis 55% MeCN in 0,1 N Et<sub>3</sub>NH<sup>+</sup>OAc<sup>-</sup>, pH 7,0, über 20 min chromatographiert. Wenn die Detritylierung durchgeführt wird, werden die Oligonukleotide durch Gelausschlußchromatographie weitergereinigt. Analytische Überprüfungen der Qualität der Oligonukleotide werden mit einer PRP-Säule (Alltech, Deerfield, IL) bei alkalischem pH und mit PAGE durchgeführt. Herstellung von 2,4,6-Trichlortriazin-abgeleiteten Oligonukleotiden: 10 bis 1.000 µg 5'-terminal aminverknüpftes Oligonukleotid werden mit einem Überschuß umkristallisierter Cyanurolsäurechloride in 10% N-Methylpyrrolidon in alkalischem Puffer (vorzugsweise pH 8,3 bis 8,5) bei 19°C bis 25°C für 30 bis 120 Minuten umgesetzt. Die endgültigen Reaktionsbedingungen bestehen aus 0,15 M Natriumborat bei pH 8,3, 2 mg/ml umkristallisiertem Cyanurolsäurechlorid und 500 µg/ml entsprechendem Oligonukleotid. Das nicht-umgesetzte Cyanurolsäurechlorid wird durch Ausschlußchromatographie auf einer G-50-Sephadex-Säule (Pharmacia, Piscataway, NJ) entfernt.

[0523] Das aktivierte gereinigte Oligonukleotid wird dann mit einem 100-fachen molaren Überschuß Cystamin in 0,15 M Natriumborat bei pH 8,3 für 1 Stunde bei Raumtemperatur umgesetzt. Das nicht-umgesetzte Cystamin wird durch Ausschlußchromatographie auf einer G-50-Sephadex-Säule entfernt. Die abgeleiteten ODNs werden dann mit aminreaktiven Fluorochromen umgesetzt. Die abgeleitete ODN-Zubereitung wird in 3 Portionen aufgeteilt und jede Portion wird mit entweder (a) einem 20-fachen molaren Überschuß von Texasrot-Sulfonylchlorid (Molecular Probes, Eugene, OR), (b) einem 20-fachen molaren Überschuß von Lissamin-Sulfonylchlorid (Molecular Probes, Eugene, OR), (c) einem 20-fachen molaren Überschuß von Fluoresceinisothiocyanat umgesetzt. Die endgültigen Reaktionsbedingungen bestehen aus 0,15 M Natriumborat bei pH 8,3 für 1 Stunde bei Raumtemperatur. Die nicht-umgesetzten Fluorochrome werden durch Ausschlußchromatographie auf einer G-50-Sephadex-Säule entfernt.

[0524] Um das Fluorochrom vom Nukleotid abzuspalten, werden die ODNs auf  $1 \times 10^{-5}$  molar eingestellt und dann Verdünnungen (12,3-fache Verdünnungen) in TE hergestellt (TE ist 0,01 M Tris, pH 7,0, 5 mM EDTA). Zu 100 µl Volumina von ODNs werden 25 µl 0,01 M Dithiothreitol (DTT) zugegeben. Zu einem identischen Satz von Kontrollzusammensetzungen wird kein DTT zugegeben. Die Mischung wird für 15 Minuten bei Raumtemperatur inkubiert. Fluoreszenz wird in einer schwarzen Mikrotiterplatte gemessen. Die Lösung wird aus den Inkubationsröhren (150 µl) entfernt und in eine schwarze Mikrotiterplatte (Dynatek Laboratories, Chantilly, VA) gegeben. Die Platten werden dann direkt unter Verwendung eines Fluoroskan-II-Fluorometers (Flow Laboratories, McLean, VA) unter Verwendung einer Anregungswellenlänge von 495 nm und Überwachung der Emission bei 520 nm für Fluorescein, unter Verwendung einer Anregungswellenlänge von 591 nm und Überwachung der Emission bei 612 nm für Texasrot und unter Verwendung einer Anregungswellenlänge von 570 nm und Überwachung der Emission bei 590 nm für Lissamin abgelesen, wobei die Ergebnisse in Tabelle I angegeben sind.

Tabelle I

Mole an Fluorochrom	RFU Nicht-gespalten	RFU gespalten	RFU frei
$1,0 \times 10^5 \text{ M}$	6,4	1200	1345
$3,3 \times 10^6 \text{ M}$	2,4	451	456
$1,1 \times 10^6 \text{ M}$	0,9	135	130
$3,7 \times 10^7 \text{ M}$	0,3	44	48
$1,2 \times 10^7 \text{ M}$	0,12	15,3	16,0
$4,1 \times 10^7 \text{ M}$	0,14	4,9	5,1
$1,4 \times 10^8 \text{ M}$	0,13	2,5	2,8
$4,5 \times 10^9 \text{ M}$	0,12	0,8	0,9

[0525] Die Daten zeigen, daß es einen etwa 200-fachen Anstieg der relativen Fluoreszenz gibt, wenn das Fluorochrom vom ODN abgespalten wird.



## BEISPIEL 4

## HERSTELLUNG VON MARKIERTEN M13-SEQUENZPRIMERN UND NACHWEIS DER ABSPALTUNG VON MARKIERUNGEN

[0526] Herstellung von 2,4,6-Trichlortriazin-abgeleiteten Oligonukleotiden: 1.000 µg 5'-terminales aminverknüpftes Oligonukleotid (5'-Hexylamin-TGTAAAACGACGGCCAGT-3") (Seq. ID Nr. 1) werden mit einem Überschuß umkristallisierten Cyanursäurechlorids in 10% N-Methylpyrrolidon in alkalischem Puffer (vorzugsweise pH 8,3 bis 8,5) bei 19 bis 25°C für 30 bis 120 Minuten umgesetzt. Die endgültigen Reaktionsbedingungen bestehen aus 0,15 M Natriumborat bei pH 8,3, 2 mg/ml umkristallisiertem Cyanursäurechlorid und 500 µg/ml entsprechendem Oligonukleotid. Das nicht-umgesetzte Cyanursäurechlorid wird durch Ausschlußchromatographie auf einer G-50-Sephadex-Säule entfernt.

[0527] Das aktivierte gereinigte Oligonukleotid wird anschließend mit einem 100-fachen molaren Überschuß Cystamin in 0,15 M Natriumborat bei pH 8,3 für 1 Stunde bei Raumtemperatur umgesetzt. Das nicht-umgesetzte Cystamin wird durch Ausschlußchromatographie auf einer G-50-Sephadex-Säule entfernt. Die abgeleiteten ODNs werden anschließend mit einer Vielzahl von Amidinen umgesetzt. Die abgeleitete ODN-Zubereitung wird in 12 Portionen aufgeteilt und je Portion wird (25-molarer Überschuß) mit den Pentafluorphenylestern von: (1) 4-Methoxybenzoesäure, (2) 4-Fluorbenzoesäure, (3) Tolylsäure, (4) Benzoesäure, (5) Indol-3-essigsäure, (6) 2,6-Difluorbenzoesäure, (7) Nicotinsäure-N-oxid, (8) 2-Nitrobenzoesäure, (9) 5-Acetylsalicylsäure, (10) 4-Ethoxybenzoesäure, (11) Zimtsäure oder (12) 3-Aminorricotinsäure umgesetzt. Die Reaktion läuft für 2 Stunden bei 37°C in 0,2 M NaBorat pH 8,3. Die abgeleiteten ODNs werden mit Gelausschlußchromatographie auf G-50-Sephadex gereinigt.

[0528] Um die Markierung vom Oligonukleotid abzuspalten, werden die ODNs auf  $1 \times 10^{-5}$  molar eingestellt und anschließend werden Verdünnungen (12,3-fache Verdünnungen) in TE hergestellt (TE ist 0,01 M Tris, pH 7,0, 5 mM EDTA) mit 50% EtOH (v/v). Zu 100 µl-Volumina ODNs werden 25 µl 0,01 M Dithiothreitol (DTT) zugegeben. Zu einem identischen Satz von Kontrollzusammensetzungen wird kein DTT zugegeben. Die Inkubation erfolgt für 30 Minuten bei Raumtemperatur. NaCl wird anschließend auf 0,1 M zugegeben und 2 Volumina EtOH werden zugegeben, um die ODNs auszufällen. Die ODNs werden aus Lösung durch Zentrifugation bei  $14.000 \times G$  bei 4°C für 15 Minuten entfernt. Die Überstände werden aufbewahrt, vollständig getrocknet. Der Pellet wird anschließend in 25 µl McOH gelöst. Der Pellet wird anschließend mit Massenspektrometrie auf das Vorhandensein von Markierungen getestet.

[0529] Der Massenspektrometer, der bei diesen Arbeiten verwendet wurde, ist ein Fouriertransformati-ons-Massenspektrometer (FTMS) mit externer Ionenquelle. Proben, die für MALDI-Analyse hergestellt wurden, werden auf der Spitze einer direkten Sonde abgeschieden und in die Ionenquelle eingeführt. Wenn die Probe mit einem Laserpuls bestrahlt wird, werden Ionen aus der Quelle extrahiert und in eine lange Quadrapol-Ionenführung geleitet, die sie fokussiert und zu einer FTMS-Analysatorzelle transportiert, die innerhalb der Bohrung eines supraleitfähigen Magneten angeordnet ist.

[0530] Die Spektren liefern die folgende Information. Peaks die in der Intensität von 25 bis 100 relativen Intensitätseinheiten variieren, bei den folgenden relativen Molekülmassen: (1) 212,1 amu, was 4-Methoxybenzoesäure-Derivat anzeigt, (2) 200,1, was 4-Fluorbenzoesäure-Derivat anzeigt, (3) 196,1 amu, was Tolylsäure-Derivat anzeigt, (4) 182,1 amu, was Benzoesäure-Derivat anzeigt, (5) 235,2 amu, was Indol-3-essigsäure-Derivat anzeigt, (6) 218,1 amu, was 2,6-Difluorbenzoesäure-Derivat anzeigt, (7) 199,1 amu, was Nikotinsäure-N-oxid-Derivat anzeigt, (8) 227,1 amu, was 2-Nitrobenzamid anzeigt, (9) 179,18 amu, was 5-Acetylsalicylsäure-Derivat anzeigt, (10) 226,1 amu, was 4-Ethoxybenzoesäure-Derivat anzeigt, (11) 209,1 amu, was Zimtsäure-Derivat anzeigt, (12) 198,1 amu, was 3-Aminorricotinsäure-Derivat hinweist.

[0531] Die Ergebnisse zeigten, daß die MW-Identifikatoren von den Primern abgespalten werden und mit Massenspektrometrie nachweisbar sind.

## BEISPIEL 5

HERSTELLUNG EINES SATZES VON VERBINDUNGEN DER FORMEL  $R_{1-36}$ -LYS( $\epsilon$ INIP)-ANP-TFP

[0532] Fig. 3 veranschaulicht die parallele Synthese eines Satzes von 36 T-L-X-Verbindungen ( $X = L_n$ ), wobei  $L_n$  ein aktivierter Ester ist (genauer Tetrafluorphenylester),  $L^2$  eine ortho-Nitrobenzylamin-Gruppe ist, wobei  $L^3$  eine Methylengruppe ist, die  $L_n$  und  $L^2$  verknüpft, T eine modulare Struktur hat, wobei die Carbonsäuregruppe von Lysin mit dem Stickstoffatom der  $L^2$ -Benzylamin-Gruppe verknüpft worden ist, um eine Amidbindung zu bilden, und eine Komponente  $R_{1-36}$  mit variabler Masse (wobei die R-Gruppen  $T^2$  entsprechen, wie hierin definiert, und über jede der spezifischen Carbonsäuren, die hierin aufgelistet sind, eingeführt werden können) durch die  $\alpha$ -Aminogruppe des Lysins gebunden ist, während eine Massenseptrometrieempfindlichkeitsverstärkergruppe (eingeführt über N-Methylisopropionylsäure) durch die  $\epsilon$ -Aminogruppe des Lysins gebunden ist.

[0533] Bezug nehmend auf **Fig. 3**:

[0534] Schritt A. NovaSyn-HMP-Harz (erhältlich von NovaBiochem; 1 Äq.) wird mit DMF im Sammelgefäß des ACT357 suspendiert. Verbindung I (ANP, erhältlich von ACT; 3 Äq.), HATU (3 Äq.) und NMM (7,5 Äq.) in DMF werden zugegeben und das Sammelgefäß für 1 h geschüttelt. Das Lösungsmittel wird entfernt und das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Die Kopplung von I an das Harz und die Waschschrte werden wiederholt, um Verbindung II zu ergeben.

[0535] Schritt B. Das Harz (Verbindung II) wird mit 25% Piperidin in DMF vermischt und für 5 min geschüttelt. Das Harz wird filtriert, anschließend mit 25% Piperidin in DMF vermischt und für 10 min geschüttelt. Das Lösungsmittel wird entfernt, das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen und direkt in Schritt C verwendet.

[0536] Schritt C. Das Harz ohne Schutzgruppe aus Schritt B wird in DMF suspendiert und dazu wird eine Aminosäure aus Fmoc-Schutzgruppe hinzugegeben, die eine geschützte Aminfunktionalität in ihrer Seitenkette enthält (Fmoc-Lysin(Aloc)-OH, erhältlich von PerSeptive Biosystems; 3 Äq.), HATU (3 Äq.) und NMM (7,5 Äq.) in DMF. Das Gefäß wird für 1 h geschüttelt. Das Lösungsmittel wird entfernt und das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Die Kopplung von Fmoc-Lys(Aloc)-OH an das Harz und die Waschschrte werden wiederholt, um Verbindung IV zu ergeben.

[0537] Schritt D. Das Harz (Verbindung IV) wird mit  $\text{CH}_2\text{Cl}_2$  (2X) gewaschen und anschließend in einer Lösung von  $(\text{PPh}_3)_4\text{Pd}(0)$  (0,3 Äq.) und  $\text{PhSiH}_3$  (10 Äq.) in  $\text{CH}_2\text{Cl}_2$  suspendiert. Die Mischung wird für 1 h geschüttelt. Das Lösungsmittel wird entfernt und das Harz mit  $\text{CH}_2\text{Cl}_2$  gewaschen. Der Palladiumschritt wird wiederholt. Das Lösungsmittel wird entfernt und das Harz mit  $\text{CH}_2\text{Cl}_2$  (2X), N,N-Düsoopropylethylammoniumdiethyldithiocarbamat in DMF (2X), DMF (2X) gewaschen, um Verbindung V zu ergeben.

[0538] Schritt E. Das Harz ohne Schutzgruppe aus Schritt D wird mit N-Methylisonipecotinsäure gekoppelt, wie in Schritt C beschrieben, um Verbindung VI zu ergeben.

[0539] Schritt F. Das Fmoc-geschützte Harz VI wird gleichmäßig durch den ACT357 aus dem Sammelgefäß in 36 Reaktionsgefäße aufgeteilt, um die Verbindungen  $\text{VI}_{1-36}$  zu ergeben.

[0540] Schritt G. Das Harz (Verbindungen  $\text{VI}_{1-36}$ ) wird mit Piperidin behandelt, wie beschrieben in Schritt B, um die Fmoc-Gruppe zu entfernen.

[0541] Schritt H. Die 36 Aliquote des Harzes ohne Schutzgruppe aus Schritt G werden in DMF suspendiert. Zu jedem Reaktionsgefäß werden die entsprechende Carbonsäure ( $\text{R}_{1-36}\text{CO}_2\text{H}$ ; 3 Äq.), HATU (3 Äq.) und NMM (7,5 Äq.) in DMF zugegeben. Die Gefäße werden für 1 h geschüttelt. Das Lösungsmittel wird entfernt und die Harz-Aliquote mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Die Kopplung von  $\text{R}_{1-36}\text{CO}_2\text{H}$  an die Harz-Aliquote und die Waschschrte werden wiederholt, um die Verbindungen  $\text{VIII}_{1-36}$  zu ergeben.

[0542] Schritt I. Die Harz-Aliquote (Verbindungen  $\text{VIII}_{1-36}$ ) werden mit  $\text{CH}_2\text{Cl}_2$  (3X) gewaschen. Zu jedem der Reaktionsgefäße wird TFA :  $\text{H}_2\text{O}$  :  $\text{CH}_2\text{Cl}_2$  im Verhältnis 90 : 5 : 5 zugegeben und die Gefäße für 120 min geschüttelt. Das Lösungsmittel wird aus den Reaktionsgefäßen in einzelne Röhren filtriert. Die Harz-Aliquote werden mit  $\text{CH}_2\text{Cl}_2$  (2X) und MeOH (2X) gewaschen und die Filtrate in die einzelnen Röhren hinein zusammengeführt. Die einzelnen Röhren werden in vacuo eingedampft, was die Verbindungen  $\text{IX}_{1-36}$  liefert.

[0543] Schritt J. Jede der freien Carbonsäure  $\text{IX}_{1-36}$  wird in DMF gelöst. Zu jeder Lösung wird Pyridin (1,05 Äq.) zugegeben, gefolgt von Tetrafluorphenyltrifluoracetat (1,1 Äq.). Die Mischungen werden für 45 min bei Raumtemperatur gerührt. Die Lösungen werden mit EtOAc verdünnt, mit 5% wäßrigem  $\text{NaHCO}_3$  (3X) gewaschen, über  $\text{Na}_2\text{SO}_4$  getrocknet, filtriert und in vacuo eingedampft, was die Verbindungen  $\text{X}_{1-36}$  liefert.

## BEISPIEL 6

### HERSTELLUNG EINES SATZES VON VERBINDUNGEN DER FORMEL $\text{R}_{1-36}\text{-LYS}(\epsilon\text{-iNIP})\text{-NBA-TFP}$

[0544] **Fig. 4** veranschaulicht die parallele Synthese eines Satzes von 36 T-L-X-Verbindungen ( $\text{X} = \text{L}_n$ ), wobei  $\text{L}_n$  ein aktivierter Ester ist (genauer Tetrafluorphenylester),  $\text{L}^2$  eine ortho-Nitrobenzylamin-Gruppe ist, wobei  $\text{L}^3$  eine direkte Bindung zwischen  $\text{L}_n$  und  $\text{L}^2$  ist, wobei  $\text{L}_n$  direkt mit dem aromatischen Ring der  $\text{L}^2$ -Gruppe verknüpft ist, T eine modulare Struktur hat, wobei die Carbonsäuregruppe von Lysin mit dem Stickstoffatom der  $\text{L}^2$ -Benzylamin-Gruppe verknüpft ist, um eine Amidbindung zu bilden, und eine Komponente  $\text{R}_{1-36}$  mit variabler Masse (wobei diese R-Gruppen  $\text{T}^2$  entsprechen, wie hierin definiert, und über jede der spezifischen Carbonsäuren, die hierin aufgelistet ist, eingeführt werden können) durch die  $\alpha$ -Aminogruppe des Lysins gebunden ist, während eine Massenspektrometrieverstärkergruppe (eingeführt über N-Methylisonipecotinsäure) durch die  $\epsilon$ -Aminogruppe des Lysins gebunden ist.

[0545] Bezug nehmend auf **Fig. 4**

[0546] Schritt A. NovaSyn-HMP-Harz wird mit Verbindung I (NBA hergestellt gemäß dem Verfahren von Brown et al., Molecular Diversity, 1, 4 (1995)) gemäß dem in Schritt A von Beispiel 5 beschriebenen Verfahren gekoppelt, um Verbindung II zu ergeben.

[0547] Schritte B–J. Das Harz (Verbindung II) wird behandelt, wie beschrieben in den Schritten B–J von Bei-

spiel 5, um die Verbindungen  $X_{1-36}$  zu ergeben.

## BEISPIEL 7

HERSTELLUNG EINES SATZES VON VERBINDUNGEN DER FORMEL  $\text{iNIP-LYS}(\epsilon\text{-R}_{1-36})\text{-ANP-TFP}$ 

[0548] **Fig. 5** veranschaulicht die parallele Synthese eines Satzes von 36 T-L-X-Verbindungen ( $X = L_n$ ), wobei  $L_n$  ein aktivierter Ester ist (genauer Tetrafluorphenylester),  $L^2$  eine ortho-Nitrobenzylamin-Gruppe ist, wobei  $L^3$  eine Methylengruppe ist, die  $L_n$  und  $L^2$  verknüpft, T eine modulare Struktur hat, wobei die Carbonsäuregruppe von Lysin mit dem Stickstoffatom der  $L^2$ -Benzylamin-Gruppe verknüpft worden ist, um eine Amidbindung zu bilden, und eine Komponente  $R_{1-36}$  mit variabler Masse (wobei diese R-Gruppen  $T^2$  entsprechen, wie hierin definiert, und über jede der spezifischen Carbonsäuren, die hierin aufgelistet sind, eingeführt werden können) durch die  $\epsilon$ -Aminogruppe des Lysins gebunden ist, während eine Massenspektrometrieempfindlichkeitsverstärkergruppe (eingeführt über N-Methylisonipicotinsäure) durch die  $\alpha$ -Aminogruppe des Lysins gebunden ist.

[0549] Bezug nehmend auf **Fig. 5**:

[0550] Schritte A–C. Dieselben wie in Beispiel 5.

[0551] Schritt D. Das Harz (Verbindung IV) wird mit Piperidin behandelt, wie in Schritt B von Beispiel 5 beschrieben, um die Fmoc-Gruppe zu entfernen.

[0552] Schritt E. Das  $\alpha$ -Amin ohne Schutzgruppe auf dem Harz in Schritt D wird mit N-Methylisonipicotinsäure gekoppelt, wie beschrieben in Schritt C von Beispiel 5, um Verbindung V zu ergeben.

[0553] Schritt F. Derselbe wie in Beispiel 5.

[0554] Schritt G. Das Harz (Verbindungen  $VI_{1-36}$ ) wird behandelt mit Palladium, wie beschrieben in Schritt D von Beispiel 5, um die Aloc-Gruppe zu entfernen.

[0555] Schritte H–J. Die Verbindungen  $X_{1-36}$  werden auf dieselbe Art und Weise hergestellt wie in Beispiel 5.

## BEISPIEL 8

HERSTELLUNG EINES SATZES VON VERBINDUNGEN DER FORMEL  $R_{1-36}\text{-GLU}(\gamma\text{-DIAEA})\text{-ANP-TFP}$ 

[0556] **Fig. 6** veranschaulicht die parallele Synthese eines Satzes von 36 T-L-X-Verbindungen ( $X = L_n$ ), wobei  $L_n$  ein aktivierter Ester ist (genauer Tetrafluorphenylester),  $L^2$  eine ortho-Nitrobenzylamin-Gruppe ist, wobei  $L^3$  eine Methylengruppe ist, die  $L_n$  und  $L^2$  verknüpft, T eine modulare Struktur hat, wobei die  $\alpha$ -Carbonsäuregruppe von Glutaminsäure mit dem Stickstoffatom der  $L^2$ -Benzylamin-Gruppe verknüpft worden ist, um eine Amidbindung zu bilden, und eine Komponente  $R_{1-36}$  mit variabler Masse (wobei diese R-Gruppen  $T^2$  entsprechen, wie hierin definiert, und über eine der spezifischen Carbonsäuren, die hierin aufgelistet sind, eingeführt werden können) durch die  $\alpha$ -Aminogruppe der Glutaminsäure gebunden ist, während eine Massenspektrometrieempfindlichkeitsverstärkergruppe (eingeführt über 2-(Diisopropylamino)ethylamin) durch die  $\gamma$ -Carbonsäure der Glutaminsäure gebunden ist.

[0557] Bezug nehmend auf **Fig. 6**:

[0558] Schritte A–B. Dieselben wie in Beispiel 5.

[0559] Schritt C. Das Harz ohne Schutzgruppe (Verbindung III) wird an Fmoc-Glu-(OAI)-OH unter Verwendung des Kopplungsverfahrens, das in Schritt C von Beispiel 5 beschrieben ist, gekoppelt, um Verbindung IV zu ergeben.

[0560] Schritt D. Der Allylester auf dem Harz (Verbindung IV) wird mit  $\text{CH}_2\text{Cl}_2$  (2X) gewaschen und mit einer Lösung von  $(\text{PPh}_3)_4\text{Pd}(0)$  (0,3 Äq.) und N-Methylanilin (3 Äq.) in  $\text{CH}_2\text{Cl}_2$  vermischt. Die Mischung wird für 1 h geschüttelt. Das Lösungsmittel wird entfernt und das Harz wird mit  $\text{CH}_2\text{Cl}_2$  (2X) gewaschen. Der Palladiumschritt wird wiederholt. Das Lösungsmittel wird entfernt und das Harz wird mit  $\text{CH}_2\text{Cl}_2$  (2X), N,N-Diisopropylethylammoniumdiethyldithiocarbamat in DMF (2X), DMF (2X) gewaschen, um Verbindung V zu ergeben.

[0561] Schritt E. Das Harz ohne Schutzgruppe aus Schritt D wird in DMF suspendiert und durch Vermischen mit HATU (3 Äq.) und NMM (7,5 Äq.) aktiviert. Die Gefäße werden für 15 min geschüttelt. Das Lösungsmittel wird entfernt und das Harz mit NMP (1X) gewaschen. Das Harz wird mit 2-(Diisopropylamino)ethylamin (3 Äq.) und NMM (7,5 Äq.) vermischt. Die Gefäße werden für 1 Stunde geschüttelt. Die Kopplung von 2-(Diisopropylamino)ethylamin an das Harz und die Waschschrte werden wiederholt, um Verbindung VI zu ergeben.

[0562] Schritte F–J. Dieselben wie in Beispiel 5.

## BEISPIEL 9

HERSTELLUNG EINES SATZES VON VERBINDUNGEN DER FORMEL  $R_{1-36}$ -LYS( $\epsilon$ -iN-IP)-ANP-LYS( $\epsilon$ -NH<sub>2</sub>)-NH<sub>2</sub>

[0563] **Fig. 7** veranschaulicht die parallele Synthese eines Satzes von 36 T-L-X-Verbindungen ( $X = L_n$ ), wobei  $L_n$  ein Amin ist (genauer die  $\epsilon$ -Aminogruppe einer von Lysin abgeleiteten Einheit),  $L^2$  eine ortho-Nitrobenzylamin-Gruppe ist, wobei  $L^3$  eine Carboxamido-substituierte Alkylenaminoacylalkylen-Gruppe ist, die  $L_n$  und  $L^2$  verknüpft, T eine modulare Struktur hat, wobei die Carbonsäuregruppe von Lysin mit dem Stickstoffatom der  $L^2$ -Benzylamin-Gruppe verknüpft worden ist, um eine Amidbindung zu bilden, und eine Komponente  $R_{1-36}$  mit variabler Masse (wobei diese R-Gruppen  $T^2$  entsprechen, wie hierin definiert, und über eine der spezifischen Carbonsäuren, die hierin aufgelistet sind, eingeführt werden können) durch die  $\alpha$ -Aminogruppe des Lysins gebunden ist, während eine Massenspektrometrieempfindlichkeitsverstärkergruppe (eingeführt über N-Methylisonipecotinsäure) durch die  $\epsilon$ -Aminogruppe des Lysins gebunden ist.

[0564] Bezug nehmend auf **Fig. 7**:

[0565] Schritt A. Fmoc-Lys(Boc)-SRAM-Harz (erhältlich von ACT; Verbindung I) wird mit 25% Piperidin in DMF vermischt und für 5 min geschüttelt. Das Harz wird filtriert, dann mit 25% Piperidin in DMF vermischt und für 10 min geschüttelt. Das Lösungsmittel wird entfernt, das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen und direkt in Schritt B verwendet.

[0566] Schritt B. Das Harz (Verbindung II), ANP (erhältlich von ACT; 3 Äq.), HATU (3 Äq.) und NMM (7,5 Äq.) in DMF werden zugegeben und das Sammelgefäß für 1 h geschüttelt. Das Lösungsmittel wird entfernt und das Harz mit NMP (2X), MeOH (2X) und DMF (2X) gewaschen. Die Kopplung von I an das Harz und die Waschschritte werden wiederholt, um Verbindung III zu ergeben.

[0567] Schritte C–J. Das Harz (Verbindung III) wird behandelt wie in den Schritten B–I in Beispiel 5, um die Verbindungen  $X_{1-36}$  zu ergeben.

## BEISPIEL 10

HERSTELLUNG EINES SATZES VON VERBINDUNGEN DER FORMEL  $R_{1-36}$ -LYS( $\epsilon$ -TFA)-LYS( $\epsilon$ -INP)-ANP-TFP

[0568] **Fig. 8** veranschaulicht die parallele Synthese eines Satzes von 36 T-L-X-Verbindungen ( $X = L_n$ ), wobei  $L_n$  ein aktivierter Ester ist (genauer Tetrafluorphenylester),  $L^2$  eine ortho-Nitrobenzylamin-Gruppe ist, wobei  $L^3$  eine Methylengruppe ist, die  $L_n$  und  $L^2$  verknüpft, T eine modulare Struktur hat, wobei die Carbonsäuregruppe eines ersten Lysins mit dem Stickstoffatom der  $L^2$ -Benzylamin-Gruppe verknüpft worden ist, um eine Amidbindung zu bilden, eine Massenspektrometrieempfindlichkeitsverstärkergruppe (eingeführt über N-Methylisonipecotinsäure) durch die  $\epsilon$ -Aminogruppe des ersten Lysins gebunden ist, ein zweites Lysinmolekül mit dem ersten Lysin über die  $\alpha$ -Aminogruppe des ersten Lysins verknüpft worden ist, eine Gruppe zur Einstellung der relativen Molekülmasse (mit einer Trifluoracetylstruktur) durch die  $\epsilon$ -Aminogruppe des zweiten Lysins gebunden ist und eine Komponente  $R_{1-36}$  mit variabler Masse (wobei diese R-Gruppen  $T^2$  entsprechen, wie hierin definiert, und über eine der spezifischen Carbonsäuren, die hierin aufgelistet sind, eingeführt werden können) durch die  $\alpha$ -Aminogruppe des zweiten Lysins gebunden ist.

[0569] Bezug nehmend auf **Fig. 8**:

[0570] Schritte A–E. Diese Schritte sind identisch mit den Schritten A–E in Beispiel 5.

[0571] Schritt F. Das Harz (Verbindung VI) wird mit Piperidin behandelt wie beschrieben in Schritt B in Beispiel 5, um die Fmoc-Gruppe zu entfernen.

[0572] Schritt G. Das Harz ohne Schutzgruppe (Verbindung VII) wird an Fmoc-Lys(Tfa)-OH unter Verwendung des Kopplungsverfahrens, das in Schritt C von Beispiel 5 beschrieben ist, gekoppelt, um Verbindung VIII zu ergeben.

[0573] Schritte H–K. Das Harz (Verbindung VIII) wird behandelt, wie in den Schritten F–J in Beispiel 5, um die Verbindungen  $XI_{1-36}$  zu ergeben.

## BEISPIEL 11

HERSTELLUNG EINES SATZES VON VERBINDUNGEN DER FORMEL  $R_{1-36}$ -LYS( $\epsilon$ -iNIP)-ANP-5'-AH-ODN

[0574] **Fig. 9** veranschaulicht die parallele Synthese eines Satzes von 36 T-L-X-Verbindungen ( $X = MOI$ , wobei MOI ein Nukleinsäurefragment ist, ODN), abgeleitet von den Estern von Beispiel 5 (dasselbe Verfahren könnte bei anderen T-L-X-Verbindungen verwendet werden, in denen X ein aktivierter Ester ist). Das MOI ist an T-L durch das 5'-Ende des MOIs konjugiert, über eine Phosphodiester-Alkylenamin-Gruppe.

[0575] Bezug nehmend auf **Fig. 9**:

[0576] Schritt A. Die Verbindungen XII<sub>1-36</sub> werden hergestellt gemäß einem modifizierten Biotinylierungsverfahren in Van Ness et al., Nucleic Acids Res., 19, 3345 (1991). Zu einer Lösung eines der 5'-Amino-hexyl-Oligonukleotide (Verbindungen XI<sub>1-36</sub>, 1 mg) in 200 mM Natriumborat (pH 8,3, 250 ml) wird einer der Tetrafluorphenylester zugegeben (Verbindungen X<sub>1-36</sub> aus Beispiel A, 100-facher molarer Überschuß in 250 ml NMP). Die Reaktion wird über Nacht bei Umgebungstemperatur inkubiert. Die nicht-umgesetzten und hydrolysierten Tetrafluorphenylester werden von den Verbindungen XII<sub>1-36</sub> durch Sephadex-G-50-Chromatographie getrennt.

#### BEISPIEL 12

##### HERSTELLUNG EINES SATZES VON VERBINDUNGEN DER FORMEL R<sub>1-36</sub>-LYS(ε-iN-IP)-ANP-LYS(ε-(MCT-5'-AH-ODN))-NH<sub>2</sub>

[0577] **Fig. 10** veranschaulicht die parallele Synthese eines Satzes von 36 T-L-X-Verbindungen (X = MOI, wobei MOI ein Nukleinsäurefragment ist, ODN), abgeleitet von den Aminen von Beispiel 11 (dasselbe Verfahren könnte bei anderen T-L-X-Verbindungen verwendet werden, in denen X ein Amin ist). Das MOI ist an T-L durch das 5'-Ende des MOIs konjugiert, über eine Phosphodiester-Alkylamin-Gruppe.

[0578] Bezug nehmend auf **Fig. 10**:

[0579] Schritt A. Die 5'-[6-(4,6-Dichlor-1,3,5-triazin-2-ylamino)hexyl]-Oligonukleotide XII<sub>1-36</sub> werden hergestellt wie beschrieben in Van Ness et al., Nucleic Acids Res., 19, 3345 (1991).

[0580] Schritt B. Zu einer Lösung eines der 5'-[6-(4,6-Dichlor-1,3,5-triazin-2-ylamino)hexyl]-Oligonukleotide (Verbindungen XII<sub>1-36</sub>) in einer Konzentration von 1 mg/ml in 100 mM Natriumborat (pH 8,3) wurde ein 100-facher molarer Überschuß eines primäre Amins zugegeben, das ausgewählt ist aus R<sub>1-36</sub>-Lys(ε-iN-IP)-ANP-Lys(ε-NH<sub>2</sub>)-NH<sub>2</sub> (Verbindungen X<sub>1-36</sub> aus Beispiel 11). Die Lösung wird über Nacht bei Umgebungstemperatur gemischt. Das nicht-umgesetzte Amin wird durch Ultrafiltration durch eine Membran mit einem 3.000 MG-Schnitt (Amicon, Beverly, MA) unter Verwendung von H<sub>2</sub>O als der Waschlösung (3X) entfernt. Die Verbindungen XIII<sub>1-36</sub> werden durch Verringerung des Volumens auf 100 ml isoliert.

#### BEISPIEL 13

##### DEMONSTRATION DES GLEICHZEITIGEN NACHWEISES MEHRERER MARKIERUNGEN DURCH MASSENSPEKTROMETRIE

[0581] Dieses Beispiel liefert eine Beschreibung der Fähigkeit, mehrere Verbindungen (Markierungen) durch Massenspektrometrie gleichzeitig nachzuweisen. In diesem besonderen Beispiel werden 31 Verbindungen mit einer Matrix vermischt, auf einem festen Träger abgeschieden und getrocknet und anschließend mit einem Laser desorbiert. Die resultierenden Ionen werden anschließend in einen Massenspektrometer eingeführt.

[0582] Die folgenden Verbindungen (bezogen von Aldrich, Milwaukee, WI) werden auf einer äquimolaren Basis auf eine Endkonzentration von 0,002 M (pro Verbindung) zusammengemischt: Benzamid (121,14), Nicotinamid (122,13), Pyrazinamid (123,12), 3-Amino-4-pyrazolcarbonsäure (127,10), 2-Thiophencarboxamid (127,17), 4-Aminobenzamid (135,15), Tolumid (135,17), 6-Methylnicotinamid (136,15), 3-Aminonicotinamid (137,14), Nicotinamid-N-Oxid (138,12), 3-Hydropicolinamid (138,13), 4-Fluorbenzamid (139,13), Zimtsäureamid (147,18), 4-Methoxybenzamid (151,17), 2,6-Difluorbenzamid (157,12), 4-Amino-5-imidazolcarbonsäureamid (162,58), 3,4-Pyridindicarbonsäureamid (165,16), 4-Ethoxybenzamid (165,19), 2,3-Pyrazindicarboxamid (166,14), 2-Nitrobenzamid (166,14), 3-Fluor-4-methoxybenzoesäure (170,4), Indol-3-acetamid (174,2), 5-Acetylsalicylamid (179,18), 3,5-Dimethoxybenzamid (181,19), 1-Naphthalinacetamid (185,23), 8-Chlor-3,5-diamino-2-pyrazincarboxamid (187,59), 4-Trifluormethylbenzamid (189,00), 5-Amino-5-phenyl-4-pyrazolcarboxamid (202,22), 1-Methyl-2-benzylmalonamat (207,33), 4-Amino-2,3,5,6-tetrafluorbenzamid (208,11), 2,3-Naphthalindicarbonsäure (212,22). Die Verbindungen werden in der oben beschriebenen Konzentration in DMSO gegeben. Ein µl des Materials wird dann mit Alpha-Cyano-4-hydroxymizsäure-Matrix (nach einer 1 : 10.000-Verdünnung) vermischt und auf einem festen Träger aus rostfreiem Stahl abgeschieden.

[0583] Das Material wird anschließend mit einem Laser unter Verwendung des Protein-TOF-Massenspektrometers (Bruker, Manning Park, MA) desorbiert und die resultierenden Ionen werden in sowohl dem linearen als auch dem Reflektoren-Betriebsmodus gemessen. Die folgenden m/z-Werte werden beobachtet (**Fig. 11**):

121,1	Benzamid (121,14)
122,1	Nicotinamid (122,13)
123,1	Pyrazinamid (123,12)
124,1	
125,2	
127,3	3-Amino-4-pyrazolcarbonsäure (127,10)
127,2	2-Thiophencarboxamid (127,17)
135,1	4-Aminobenzamid (135,15)
135,1	Tolumid (135,17)
136,2	6-Methylnicotinamid (136,15)
137,1	3-Aminonicotinamid (137,14)

138,2 →	Nicotinamid-N-oxid (138,12)
138,2 →	3-Hydropicolinamid (138,13)
139,2 →	4-Fluorbenzamid (139,13)
140,2	
147,3 →	Zimtsäureamid (147,18)
148,2	
149,2	
	4-Methoxybenzamid (151,17)
152,2	
	2,6-Difluorbenzamid (157,12)
158,3	
	4-Amino-5-imidazolcarbonsäureamid (162,58)
163,3	
165,2 →	3,4-Pyridindicarbonsäureamid (165,16)
165,2 →	4-Ethoxybenzamid (165,19)
166,2 →	2,3-Pyrazindicarboxamid (166,14)
166,2 →	2-Nitrobenzamid (166,14)
	3-Fluor-4-methoxybenzoesäure (170,4)
171,1	
172,2	
173,4	
	Indol-3-acetamid (174,2)
178,3	
179,3 →	5-Acetylsalicylamid (179,18)
181,2 →	3,5-Dimethoxybenzamid (181,19)
182,2 →	
	1-Naphthalinacetamid (185,23)
186,2	
	8-Chlor-3,5-diamino-2-pyrazincarbonsäureamid (187,59)
188,2	
189,2 →	4-Trifluormethylbenzamid (189,00)

190,2

191,2

192,3

5-Amino-5-phenyl-4-pyrazolcarboxamid (202,22)

203,2

203,4

1-Methyl-2-benzylmalonamat (207,33)

4-Amino-2,3,5,6-tetrafluorbenzamid (208,11)

212,2 → 2,3-Naphthalindicarbonsäure (212,22).

219,3

221,2

228,2

234,2

237,4

241,4

[0584] Die Daten zeigen, daß 22 von 31 Verbindungen im Spektrum mit der vorhergesagten Masse erschienen, 9 von 31 Verbindungen im Spektrum mit einer  $n + H$ -Masse (1 Atommasseneinheit, amu) über der vorhergesagten Masse erschienen. Das letztere Phänomen beruht wahrscheinlich auf der Protonierung eines Amins in den Verbindungen. Daher werden 31 von 31 Verbindungen mit MALDI-Massenspektroskopie nachgewiesen. Noch wichtiger ist, daß das Beispiel zeigt, daß mehrere Markierungen gleichzeitig mit einem spektroskopischen Verfahren nachgewiesen werden können.

[0585] Die Alpha-Cyano-Matrix allein (**Fig. 11**) ergab Peaks bei 146,2, 164,1, 172,1, 173,1, 189,1, 190,1, 191,1, 192,1, 212,1, 224,1, 228,0, 234,3. Weitere identifizierte Massen im Spektrum beruhen auf Verunreinigungen in den gekauften Verbindungen, da keine Anstrengungen unternommen wurden, die Verbindungen weiter zu reinigen.

#### BEISPIEL 14

##### MIKROSATELLITEN-MARKER:PCR-AMPLIFIKATIONEN.

[0586] Die Mikrosatelliten-Marker werden amplifiziert unter Verwendung der folgenden Standard-PCR-Bedingungen. Kwz gesagt, werden PCR-Reaktionen in einem Gesamtvolumen von 50 µl durchgeführt, das 40 ng genomische DNA, 50 pmol von jedem Primer, 0,125 mM dNTPs und 1 Einheit Taq-Polymerase enthält. 1X-Amplifikationspuffer enthält 10 mM Tris-Base, pH 9, 50 mM KCl, 1,5 mM MgCl<sub>2</sub>, 0,1% Triton X-100 und 0,01% Gelatine. Die Reaktionen werden unter Verwendung eines "Heißstart"-Verfahrens durchgeführt: Taq-Polymerase wird erst nach einem ersten Denaturierungsschritt von 5 Minuten bei 96°C zugegeben. Amplifikation wird über 35 Zyklen durchgeführt: Denaturierung (94°C für 40 s) und Annealing (55°C für 30 s). Ein Verlängerungsschritt (72°C für 2 Minuten) beendet das Verfahren nach dem letzten Annealing. Da die Amplifikationsprodukte, die erhalten werden sollen, kurz sind (90 bis 350 Basenpaare lang) und das Zeitintervall, um die Temperatur von 55°C auf 94°C anzuheben (erhalten mit einer Anstiegsgeschwindigkeit von 1°C/Sekunde), lang genug ist, kann die Vervollständigung der DNA-Verlängerung ohne einen Schritt bei 72°C erreicht werden.

#### BEISPIEL 15

##### AUFRENNUNG VON DNA-FRAGMENTEN

##### Instrumentierung

[0587] Die Auftrennung von DNA-Fragmenten kann unter Verwendung eines HPLC-Systems durchgeführt



werden, das aus verschiedenen Standardkomponenten zusammengesetzt ist. Diese Komponenten schließen ein Minimum von zwei Hochdruckpumpen ein, die Lösungsmittel durch einen Hochdruckgradientenmischer, einen Injektor, einer HPLC-Säule und einen Detektor pumpen. Der Injektor ist ein automatisierter programmierbarer Autosampler, der in der Lage ist, typischerweise zwischen 80 und 100 Proben bei oder unterhalb von Raumtemperaturen zu lagern, um die Stabilität der Probenkomponenten aufrechtzuerhalten. Der Autoinjektor ist ebenfalls in der Lage,  $\mu$ l-Größen-Injektionen auf eine reproduzierbare Weise vollständig ohne Kontrolle durchzuführen. Die HPLC-Säule ist in einem geheizten Säulenkompartiment enthalten, das in der Lage ist, eine definierte Temperatur auf innerhalb von  $0,1^{\circ}\text{C}$  zu halten. Die in den Beispielen unten verwendete Säule wurde von SeraSep (San Jose, CA) unter dem Namen DNASep erworben. Die Säule ist eine  $55 \times 4,6$  mm Säule mit einem  $2,2 \mu\text{m}$  nicht-porösen Polystyroldivinylbenzol-Copolymer-Partikel, der mit C18 alkyliert ist. Das Packmaterial ist innerhalb eines pH-Bereichs von 2–12 stabil und toleriert Temperaturen so hoch wie  $70^{\circ}\text{C}$ . Der Nachweis von Analyt wurde unter der Verwendung eines Einzel- oder multiple Wellenlängen UV-Detektors oder Dioden-Arraydetektor durchgeführt.

#### Verfahren

[0588] Die in diesem Beispiel zur Abtrennung von DNA-Fragmenten angewendeten Verfahren verwenden Ionen-Paarchromatographie, eine Form von Chromatographie, in der Ionen in Lösung gepaart oder neutralisiert werden können und als ein Ionenpaar auf eine Umkehrphasensäule getrennt werden können. Der lipophile Charakter und die Konzentration des Gegenions bestimmen das Ausmaß der Retention des Analyten. Im Fall eines DNA-Moleküls paart die lipophile kationische Pufferkomponente mit anionischen Phosphatgruppen auf dem DNA-Rückgrat. Die Pufferkomponenten interagieren auch mit den Alkylgruppen der stationären Phase. Die gepaarte DNA eluiert dann gemäß ihrer Größe, während die mobile Phase fortschreitend organischer mit ansteigenden Konzentrationen von Acetonitril gemacht wird. Die Evaluierung der Geeignetheit von verschiedenen Ammoniumsalzen wurde unter der Verwendung von enzymatischem Verdau von Plasmiden oder kommerziell erhältlichen DNA-Leitern untersucht. Der Bereich von Acetonitril, der erforderlich ist um die DNA zu eluieren, sowie die Temperatur des Säulenkompartiments variierte mit jedem untersuchten Puffer.

#### Puffer

[0589] Die auf ihre Ionen-Paarungsfähigkeit hin untersuchten Puffer wurden aus Vorratslösungen hergestellt. Um die Konzentration von Ionen-Paareagenz während des Gradienten konstant zu halten, wurde das Ionen-Paareagenz zu sowohl der Wasser- und der Acetonitril-mobilen Phase hinzugefügt. Die Säule wurde mit einer neuen mobilen Phase für ungefähr 18 Stunden bei einer Flußrate von  $50 \mu\text{l}/\text{Minute}$  vor dem Versuch jeglicher Trennung äquilibriert. Sobald eine mobile Phase untersucht wurde, wurde sie entfernt und die Säule mit einer Spülung von 800 ml 0,1% Ameisensäure und 50% Acetonitril, gefolgt von einer Spülung mit 800 ml 0,1% Essigsäure und 50% Acetonitril vor der Äquilibrierung mit einer neuen mobilen Phase gereinigt.

#### A. nn-Dimethyloctylammoniumtrifluoracetat

[0590] Eine Vorratslösung von 1 molar von Dimethyloctylammoniumtrifluoracetat wurde durch Mischen einer Hälfte eines Äquivalents von Trifluoressigsäure in einem geeigneten Volumen von Wasser und langsamem Hinzufügen eines Äquivalents von nn-Dimethyloctylamin hergestellt. Der pH dieser Vorratslösung ist 7. Die Vorratslösung wurde mit einem geeigneten Volumen von Wasser oder Acetonitril auf Arbeitskonzentration verdünnt.

#### B. nn-Dimethylheptylammoniumacetat

[0591] Eine Vorratslösung von 1 molar Dimethylheptylammoniumacetat wurde durch Mischen eines Äquivalents von Eisessig in einem geeigneten Volumen von Wasser und langsamem Hinzufügen eines Äquivalents von nn-Dimethylheptylamin hergestellt. Der pH dieser Vorratslösung ist 6,6. Die Vorratslösung wurde mit einer geeigneten Menge von Wasser oder Acetonitril auf Arbeitskonzentration verdünnt.

#### C. nn-Dimethylhexylammoniumacetat

[0592] Eine Vorratslösung von 1 molar Dimethylhexylammoniumacetat wurde durch Mischen eines Äquivalents von Eisessig in einem geeigneten Volumen von Wasser und langsamem Hinzufügen eines Äquivalents von nn-Dimethylhexylamin hergestellt. Der pH dieser Vorratslösung ist 6,5. Die Vorratslösung wurde mit einem geeigneten Volumen von Wasser oder Acetonitril auf Arbeitskonzentration verdünnt.

D. nn-Dimethylbutylammoniumacetat

[0593] Eine Vorratslösung von 1 molar Dimethylbutylammoniumacetat wurde durch Mischen eines Äquivalents von Eisessig in einem geeigneten Volumen von Wasser und langsamem Hinzufügen eines Äquivalents von nn-Dimethylbutylamin hergestellt. Der pH der Vorratslösung ist 6,9. Die Vorratslösung wurde mit einem geeigneten Volumen von Wasser oder Acetonitril auf Arbeitskonzentration verdünnt.

E. nn-Dimethylisopropylammoniumacetat

[0594] Eine Vorratslösung von 1 molar Dimethylisopropylammoniumacetat wurde durch Mischen eines Äquivalents von Eisessig in einem geeigneten Volumen von Wasser und langsamem Hinzufügen eines Äquivalents von nn-Dimethylisopropylamin hergestellt. Der pH der Vorratslösung ist 6,9. Die Vorratslösung wurde mit einem geeigneten Volumen von Wasser oder Acetonitril auf Arbeitskonzentration verdünnt.

F. nn-Dimethylcyclohexylammoniumacetat

[0595] Eine Vorratslösung von 1 molar Dimethylcyclohexylammoniumacetat wurde durch Mischen eines Äquivalents von Eisessig in einem geeigneten Volumen von Wasser und langsamem Hinzufügen eines Äquivalents von nn-Dimethylcyclohexylamin hergestellt. Der pH der Vorratslösung ist 6,5. Die Vorratslösung wurde mit einem geeigneten Volumen von Wasser oder Acetonitril auf Arbeitskonzentration verdünnt.

G. Methylpiperidinacetat

[0596] Eine Vorratslösung von 1 molar Methylpiperidinacetat wurde durch Mischen eines Äquivalents von Eisessig in einem geeigneten Volumen von Wasser und langsamem Hinzufügen eines Äquivalents von 1-Methylpiperidin hergestellt. Der pH der Lösung ist 7. Die Vorratslösung wurde mit einem geeigneten Volumen von Wasser oder Acetonitril auf Arbeitskonzentration verdünnt.

H. Methylpyrrolidinacetat

[0597] Eine Vorratslösung von 1 molar Methylpyrrolidinacetat wurde durch Mischen eines Äquivalents von Eisessig in einem geeigneten Volumen von Wasser und langsamem Hinzufügen eines Äquivalents 1-Methylpyrrolidin hergestellt. Der pH der Vorratslösung ist 7. Die Vorratslösung wurde in einem geeigneten Volumen von Wasser oder Acetonitril auf Arbeitskonzentration verdünnt.

I. Triethylammoniumacetat

[0598] Eine Vorratslösung von 2 molar Triethylammoniumacetat pH 7,0 wurde von Glenn Research Sterling, Virginia, erworben. Die Vorratslösung wurde mit einem geeigneten Volumen von Wasser oder Acetonitril auf Arbeitskonzentration verdünnt.

BEISPIEL 16

DNA-FINGERPRINT

[0599] DNA-Fingerprintingadaptoen werden umfassend das Folgende hergestellt: Eine Kernsequenz und eine Enzym-spezifische Sequenz. Die Struktur des EcoRI-Adapters ist 5'-CTCGTAGACTGCGTACC (SEQ ID NO. \_\_), die Struktur des MseI-Adapters ist 5'-GACGATGAGTCCTGAG.

[0600] Adaptoen für die selten schneidenden Enzyme waren identisch zu dem EcoRI mit der Ausnahme, daß kohäsive Enden verwendet wurden. ALPH-Primer bestehen aus drei Teilen: Einer Kernsequenz, einer Enzym-spezifische Sequenz und einer selektiven Verlängerungssequenz. Die EcoRI- und MseI-Primer werden wie folgt beschrieben: EcoRI: 5'-gactgcgtaaa-aattc-NNN (SEQ ID NO. \_\_); MseI: 5'-gatgagtcctgag-taa-NNN (SEQ ID NO. \_\_).

[0601] Genomische DNA wurde für 1 Stunde bei 37°C mit 5 Einheiten EcoRI und 5 Einheiten MseI und 40 µl Volumen mit 10 mM Tris-Acetate pH 7,5, 10 mM MgAce, 50 mM KAcetate, 5 mM DTT, 50 ng/Mikroliter BSA, 5 mM DTT inkubiert. Als nächstes wurden 10 µl einer Lösung, die 5 pMol EcoRI-Adaptoen, 50 pMol MseI-Adaptoen, 1 Einheit T4-Ligase, 1 mM ATP in 10 mM Tris-Acetate pH 7,5, 10 mM MgAce, 50 mM KAcetate, 5 mM DTT, 50 ng/Mikroliter BSA enthielt, hinzugeführt und die Inkubation wurde für 3 Stunden bei 37°C fortgeführt. Die Adaptoen wurden durch Hinzufügen von äquimolaren Mengen von beiden Strängen hergestellt: Die Adaptoen waren nicht phosphoryliert. Nach der Ligation wurde das Reaktionsgemisch auf 500 µl mit 10 mM

Tris-HCl, 0,1 mM EDTA pH 8,0 verdünnt und bei – 20°C gelagert.

[0602] Genetische Fingerprintingreaktionen: Amplifikationsreaktionen werden beschrieben, die DNA-Templates für die Enzymkombination EcoRI/MseI verwenden. Genomische Fingerprints mit anderen Enzymkombinationen wurden mit geeigneten Primern durchgeführt. Die Amplifikationsreaktionen verwendeten im allgemeinen zwei Oligonukleotide, eines entsprechend den EcoRI-Enden und eines entsprechend den MseI-Enden. Einer der zwei Primer war mit dem CMST-Marker markiert, bevorzugterweise der EcoRI-Primer. Die PCRs wurden unter der Verwendung von 5 ng markiertem EcoRI-Primer, 30 ng MseI-Primer, 5 Mikroliter Templat DNA, 0,4 Einheiten Taq-Polymerase, 10 mM Tris-HCl, pH 8,3, 1,5 mM MgCl<sub>2</sub>, 50 mM KCl, 0,2 mM von dATP, dGTP, dCTP, dTTP durchgeführt. Die PCR-Reaktionen unterschieden sich in Abhängigkeit von der Art der selektiven Amplifikationsverlängerung der für die Amplifikation verwendeten DNA-Fingerprintingprimer. Die DNA-Fingerprintingreaktionen mit Primern, die zwei oder drei selektive Nukleotide aufweisen, wurden für 36 Zyklen mit dem folgenden Zyklusprofil durchgeführt: Ein 30 Sekunden DNA-Denaturierungsschritt bei 94°C, ein 30 Sekunden Annealingschritt bei 55°C und dann ein 1 Minute Verlängerungsschritt bei 72°C für 1 Minute. Die Annealingtemperatur im ersten Schritt war 65°C und wurde anschließend für jeden Zyklusschritt um 0,7°C in den nächsten 12 Zyklen verringert und wurde bei 56°C für die restlichen 23 Zyklen fortgeführt. Alle Amplifikationen wurden in einem MJ-Thermocycler (Watertown, MA) durchgeführt.

[0603] Das DNA-Fingerprinting der komplexen Genome (wie zum Beispiel bei Menschen) schloß zwei Amplifikationsschritte ein. Die Präamplifikation wurde mit zwei DNA-Fingerprintings durchgeführt, die ein einzelnes selektives Nukleotid, wie oben beschrieben, aufwiesen, mit der Ausnahme, daß 30 ng von beiden DNA-Fingerprintingprimern verwendet wurden und daß diese Primer nicht mit CMST markiert wurden, nach dem Präamplifikationsschritt wurden die Reaktionsgemische 10-fach mit 10 mM Tris-HCl, 0,1 mM EDTA, pH 8,0 verdünnt und als Template für die zweite Amplifikationsreaktion verwendet. Die zweite Amplifikationsreaktion wurde wie oben beschrieben für die DNA-Fingerprintingreaktionen mit Primern durchgeführt, die die längeren selektiven Verlängerungen aufwiesen.

[0604] Die Produkte aus den Amplifikationsreaktionen wurden durch HPLC analysiert. Die HPLC wurde unter der Verwendung von automatisierter HPLC-Instrumentierung (Rainin, Emeryville, CA oder Hewlett Packard, Palo Alto, CA) durchgeführt. Nicht-gereinigte DNA-Fingerprintingprodukte, die für 3 Minuten bei 95°C vor der Injektion in einer HPLC denaturiert wurden, wurden mit einem linearen Acetonitril (ACN, J. T. Baker, NJ) Gradienten von 1,8%/Minute bei einer Flußrate von 0,9 ml/Minute eluiert. Die Start- und Endpunkte wurden gemäß der Größe der amplifizierten Produkte eingestellt. Die für die erfolgreiche Auflösung der während der DNA-Fingerprinting-Technik erzeugten Moleküle erforderliche Temperatur war 50°C. Der Ausfluß aus der HPLC wurde dann in ein Massenspektrometer zum Nachweis der Markierungen geleitet (Hewlett Packard, Palo Alto, CA).

[0605] Die folgenden Fragmente eluierten in der präsentierten Reihenfolge (die zitierte Nummer sind die Positionen innerhalb des Lambda-Genoms, bei denen eine Spaltstelle auftrat): 47, 78, 91, 733, 1456, 2176, 3275, 3419, 4349, 444, 5268, 5709, 6076, 6184, 6551, 7024, 7949, 8062, 8200, 8461, 9079, 9253, 9692, 9952, 11083, 11116, 11518, 11584, 12619, 12967, 14108, 14892, 15628, 15968, 16034, 16295, 16859, 18869, 19137, 19482, 20800, 21226, 21441, 2635, 21702, 21903, 21948, 22724, 23048, 23084, 23111, 23206, 23297, 23285, 23479, 23498, 23555, 23693, 23887, 23979, 23987, 24073, 24102, 24751, 24987, 25170, 25255, 25353, 25437, 26104, 25578, 25746, 25968, 26133, 26426, 26451, 26483, 26523, 26585, 26651, 26666, 26679, 26693, 26763, 26810, 26984, 26993, 27038, 27092, 27203, 27317, 27683, 28456, 28569, 28922, 28972, 29374, 29981, 30822, 30620, 30639, 30722, 30735, 30756, 31169, 31747, 31808, 32194, 32218, 32641, 32704, 33222, 33351, 33688, 33736, 33748, 33801, 34202, 34366, 34406, 34590, 34618, 34684, 34735, 34753, 34831, 35062, 35269, 35534, 35541, 36275, 36282, 36303, 36430, 36492, 36531, 36543, 36604, 36736, 36757, 36879, 37032, 37442, 37766, 37783, 37882, 37916, 37994, 36164, 38287, 38412, 38834, 39168, 44972, 39607, 39835, 40127, 40506, 40560, 40881, 41017, 41423, 41652, 41715, 42317, 42631, 42651, 42673, 42814, 43410, 43492, 43507, 43528, 43593, 44424, 44538, 44596, 45151, 45788, 46033, 46408, 46556, 46804, 46843, 46853, 46896, 46952, 47256, 47274, 47287, 47430, 47576, 47699, 47799, 48059, 48125, 48227, 48359, 48378. Die durchschnittliche Fragmentlänge war ungefähr 160 nt. Die beobachteten Stellen der Spaltung waren größtenteils (>95%) mit denjenigen aus einer MseI/EcoRI Verdaukarte vorhergesagten kompatibel.

#### BEISPIEL 17

#### EINZELNUKLEOTID-VERLÄNGERUNGSSASSAYS

[0606] RNA-Präparation: Gesamt-RNA wurde isoliert und wurde aus Jurkat-Zellen (ausgehend von  $1 \times 10^9$  Zellen in exponentiellen Wachstum) unter der Verwendung eines RNA-Isolierungskits von Promega (WI) präpariert. Die RNA wurde in zwei Aliquots gelagert: 1) Vorratsaliquot in Diethylpyrocarbonat-behandeltem ddH<sub>2</sub>O wurde bei –20°C gelagert und 2) Langzeitlagerung als eine Suspension in 100% H<sub>2</sub>O.

[0607] Reverse Transkription: Poly(dT)-geprimte reverse Transkription von Gesamt-RNA wurde wie beschrie-

ben in Ausubel et al. (Ausubel et al., in Current Protocols in Molecular Biology, 1991, Greene Publishing Associates/Wiley Interscience, NY, NY.) durchgeführt mit der Ausnahme, daß die Reaktionen) auf die Verwendung von 1 µg von Anfangs-Gesamt-RNA eingestellt wurden. 20–50 Einheiten von reverser Transkriptase (Promega) wurden 10-fach in 10% Glycerin, 10 mM KP04, pH 7,4, 0,2% Triton X-100, und 2 mM DTT verdünnt und auf Eis für 30 Minuten plaziert, vor der Zugabe zu den Reaktionen. Gen-spezifische reverse Transkription für GADPH und andere Kontrollgene, wie unten beschrieben, wurden unter der Verwendung von 1 µg Gesamt Jurkat RNA durchgeführt, die in 10 mM Tris-HCl pH 8,3, 50 mM KCl, 1 mM MgCl<sub>2</sub>, 1 mM dNTPs, 2 U/µl RNasin (Gibco-BRL), 0,1 M Oligomer und 0,125 U/µl von M-MLV reverse Transkriptase (Gibco-BRL) in 20 µl Reaktionen revers transkribiert wurde. Die Reaktionen wurden bei 42°C für 15 Minuten inkubiert, bei 95°C für 5 Minuten Hitze-inaktiviert und auf 100 µl mit einem Mastermix von (10 mM Tris HCl pH 8,3, 1 mM NH<sub>4</sub>Cl, 1,5 M MgCl<sub>2</sub>, 100 mM KCl), 0,125 mM NTPs, 10 ng/ml der jeweiligen Oligonucleotidprimer und 0,75 Einheiten von Taq-Polymerase (Gibco-BRL) in Vorbereitung auf die PCR-Amplifikation verdünnt.

[0608] PCR: Die PCR für jedes Gen wurde mit genspezifischen Primern durchgeführt, die eine bekannte Intron/Exon-Grenze überspannten (siehe unten). Alle PCRs wurden in 20 µl Volumen durchgeführt, die 10 mM Tris HCl pH 8,3, 1 mM NH<sub>4</sub>Cl, 1,5 M MgCl<sub>2</sub>, 100 mM KCl), 0,125 mM NTPs, 10 ng/ml der jeweiligen Oligonukleotidprimer und 0,75 Einheiten Taq-Polymerase (Gibco-BRL) enthielten. Die Zyklusparameter waren 94°C Vor-Erhitungsschritt für 5 Minuten, gefolgt von 94°C Denaturierungsschritt für 1 Minute, 55°C Annealingschritt für 2 Minuten und ein 72°C Verlängerungsschritt für 30 Sekunden bis 1 Minute und einer finalen Verlängerung bei 72°C für 10 Minuten. Die Amplifikationen waren im allgemeinen 30–5 an der Zahl.

[0609] Reinigung von Templaten: Die PCR-Produkte wurden wie durch Zhen und Swank (Zhen und Swank, BioTechniques, 14(6): 894–898, 1993) beschrieben gelgereinigt. Die PCR-Produkte wurden auf 1% Agarosegelen getrennt, die in 0,04 M Tris-Acetat, 0,001 M EDTA (1 × TAE) Puffer laufen gelassen wurden und mit Ethidiumbromid gefärbt wurden, während sie mit einer UV Lichtquelle sichtbar gemacht wurden. Eine Vertiefung wurde gerade vor der Bunde von Interesse eingeschnitten und mit 50–200 µl von 10% PEG in 1 × TAE Puffer gefüllt. Die Elektrophorese wurde fortgeführt, bis die Bunde vollständig in die Vertiefung eingetreten war. Die Inhalte wurden dann entfernt und extrahiert mit Phenol, Chloroform extrahiert, und dann in 0,1 Volumen von 7,5 M Ammoniumacetat und 2,5 Volumen 100% EtOH gefällt. Die Proben. wurden mit 75% EtOH gewaschen und kurz bei Raumtemperatur getrocknet. Die Quantifizierung der Ausbeute wurde durch Elektrophorese eines kleinen Aliquots auf 1% Agarosegel in 1 × TBE Puffer mit Ethidiumbromidfärbung und Vergleich mit einem bekannten Standard durchgeführt.

[0610] Jede SNUPE-Reaktion wurde in einem 50 µl Volumen, das ungefähr 100 ng des amplifizierten DNA Fragments, 1 µM des SNUPE-Primers, 2 Einheiten Taq-Polymerase und 1 µl des geeigneten dNTP enthielt, durchgeführt. Alle dNTPs waren in diesem Typ Assay nicht-markiert. Der verwendete Puffer war 10 mM Tris-HCl (pH 8,3), mit 50 mM KCl, 5 mM MgCl<sub>2</sub> und 0,001% (Gew/Vol) Gelatine. Die Proben wurden einem Zyklus unterzogen, der aus einer 2-Minuten Denaturierungsperiode bei 95°C, einer 2 Minuten Annealingperiode bei 60°C und einer 2-Minuten Primerverlängerungsperiode bei 72°C bestand. Die Sequenz des SNUPE-Primers für jede Familie ist unten beschrieben.

[0611] Primerverlängerungen: Einzel-Nucleotidprimerverlängerung wurden wie in Singer-Sam et al., (Singer-Sam et al., PCR Methods und Applications 1: 160–163, 1992) durchgeführt, mit der Ausnahme, daß 1 mM Mg<sup>++</sup>, 0,1 µM Primer, und 0,05 µM jedes dNTP-Typs in jedem Reaktionstyp verwendet wurden. Nach jeder oben beschriebenen Primerverlängerung wurde ein fünftel Volumen eines Ladefarbstoffs (80% Formamid, 0,1% Bromophenolblau, 0,1% Xylencyanol, 2 mM EDTA) hinzugefügt und die gesamte Probe in 15% denaturierendem Polyacrylamidgel elektrophoriert. Die Gele wurden in 10% Glycerin, 10% Methanol, 10% Eisessig unter konstantem Schütteln fixiert, gefolgt von Waschschritten mit 10% Glycerin. Die Gele wurden dann bei 55°C für 3–5 Stunden getrocknet.

[0612] Die in diesem Experiment beschriebenen Primer sind durch Rychlik beschrieben (Rychlik, BioTechniques 18: 84–90, 1995). Die Primer können synthetisiert werden oder als Gel-Filtrationsgradprimer von Midland Certified Reagent Company (Midland, Texas) erhalten werden. Die Amplifikationen sind entweder Taq DNA-Polymerase-basiert (10 mM Tris-HCl pH 8,3, 1,5 mM MgCl<sub>2</sub>, 50 mM KCl) oder Pfu-DNA-Polymerase-basiert (20 mM Tris-HCl pH 8,3, 2,0 mM MgCl<sub>2</sub>, 10 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0,1% Triton X-100, 0,1 mg/ml Rinderserumalbumin). Die Gesamt-Nukleosidtriphosphat (NTPs) Konzentration in den Reaktionen ist 0,8 mM, die Primerkonzentration ist 200 nM (falls nicht anders angegeben) und die Templatmenge ist 0,25 ng von Bakteriophagen Lambda-DNA pro 20 µl Reaktion. Die Zyklusparameter waren 94°C Vor-Heizschritt für 5 Minuten, gefolgt von 94°C Denaturierungsschritt für 1 Minute, 55°C Annealingschritt für 2 Minuten und einem 72°C Verlängerungsschritt für 30 Sekunden bis 1 Minute und einer finalen Verlängerung bei 72°C für 10 Minuten. Die Amplifikationen waren im allgemeinen 30–45 an der Zahl.

[0613] Zwei Regionen im Bakteriophagen Lambda-Genom (GenBank Zugangsnummer # J02459) wurden als die Startstellen für die Amplifikation ausgesucht. Der 5'-Primer wies ein stabiles GC-reiches 3'-Ende auf: Der 3'-Primer ist so ausgewählt, daß sich ein 381 bp Produkt ergeben wird. Der 5'-Vorwärtsprimer ist H17: 5'-GAACGAAAACCCCCCGC (SEQ ID NO: \_). Der 3'-rückwärtige Primer ist RP 17: 5'-GATCGCCCCCAAAA-

CACATA (SEQ ID NO: \_).

[0614] Das amplifizierte Produkt wurde dann auf die Anwesenheit eines Polymorphismus an Position 31245 getestet. Der folgende Primer wurde in vier Einzel-Nukleotidverlängerungstests verwendet; SNE17: 5'-GAACGAAAACCCCCCGC (SEQ ID NO: \_). Die vier Einzel-Nukleotidverlängerungsassays wurden dann wie oben beschrieben durchgeführt. All die Reaktionen wurden dann vereinigt und 5 µl des vereinigten Materials wurde auf die HPLC-Säule (SeraSep, San Jose, CA) ohne weitere Reinigung injiziert.

[0615] Die HPLC wurde unter der Verwendung von automatisierter HPLC-Instrumentierung (Rainin, Emeryville, CA, oder Hewlett Packard, Palo Alto, CA) durchgeführt. Nicht-gereinigte SNEA-Produkte, die für 3 Minuten bei 95°C vor der Injektion in eine HPLC denaturiert wurden, wurden mit einem linearen Acetonitrilgradienten (ACN, J. T. Baker, NJ) von 1,8%/Minute bei einer Flußrate von 0,9 ml/Minute eluiert. Die Start und Endpunkte wurden gemäß der Größe des SNEA-Produkts eingestellt. Die für die erfolgreiche Auflösung der SNEA-Moleküle erforderliche Temperatur war 50°C. Der Ausfluß aus der HPLC wurde dann in ein Massenspektrometer geleitet (Hewlett Packard, Palo Alto, CA) zum Nachweis von Markierungen, wobei die Ergebnisse in Tabelle 2 angegeben sind.

TABELLE 2

Markierter Primer	ddNTP-Typ	Retentionszeit	verlängert?
SNE17-487	ddATP	2,5 Minuten	nein
SNE17-496	ddGTP	2,5 Minuten	nein
SNE17-503	ddCTP	4,6 Minuten	ja
SNE17-555	ddTTP	2,5 Minuten	nein

[0616] Die Ergebnisse zeigen daher an, daß der Massenspektrometrie-Marker (CMST) bei einer Retentionszeit von 4,6 Minuten nachgewiesen wurde, was anzeigt, daß der SNE17-Primer durch eine Base (ddCTP) verlängert wurde und daher der Polymorphismus an der Position 31245 in diesem Falle ein "G" war. Die SNE17-487, SNE17-496, und SNE17-555 markierten Primer wurden nicht verlängert und ihre Retentionszeit auf der HPLC war jeweils 2,5 Minuten.

## BEISPIEL 18

[0617] In diesem Beispiel (18) wurden alle Reaktionen in Folien-abgedeckten Flaschen durchgeführt. Die Sequenz der Reaktionen A → F, beschrieben in diesem Beispiel, ist in **Fig. 19A** und **19B** dargestellt. Die in diesem Beispiel angegebenen Verbindungsnummern beziehen sich auf Verbindungen derselben Nummern in **Fig. 19A** und **19B**.

[0618] A. Zu einer Lösung eines ANP-Linkers (Verbindung 1, 11,2 mMol) und Diisopropylethylamin (22,4 mMol) in CHCl<sub>3</sub> (60 ml) wurde Allylbromid (22,4 mMol) hinzugefügt. Das Reaktionsgemisch wurde für 3 Stunden refluxiert, für 19 Stunden bei Raumtemperatur gerührt, verdünnt mit CHCl<sub>3</sub> (200 ml) und mit 1,0 M HCl (2 × 150 ml) und H<sub>2</sub>O (2 × 150 ml) gewaschen. Die organischen Extrakte wurden getrocknet (MgSO<sub>4</sub>) und das Lösungsmittel abgedampft, um Verbindung 2 als gelben Feststoff zu ergeben.

[0619] Zu einem Gemisch von Verbindung 2 in CH<sub>2</sub>Cl<sub>2</sub> (70 ml), wurde Tris(2-aminoethyl)amin (50 ml) hinzugefügt und das Reaktionsgemisch bei Raumtemperatur für 18 Stunden gerührt. Die Reaktion wurde mit CH<sub>2</sub>Cl<sub>2</sub> (150 ml) verdünnt und mit pH 6,0 Phosphatpuffer (2 × 150 ml) gewaschen. Die organischen Extrakte wurden getrocknet (MgSO<sub>4</sub>) und das Lösungsmittel abgedampft. Der Rest wurde Säulenchromatographie (Hexan/EtOAc) unterzogen, um 1,63 g (58%) von Verbindung 3 zu ergeben: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 7,85 (dd, 2H), 7,70 (t, 1H), 7,43 (t, 1H), 5,85 (m, 1H); 5,20 (q, 2H), 4,58 (q, 1H), 4,50 (d, 2H), 2,70 (m, 2H), 2,20 (br s, 2H).

[0620] B. Zu einer Lösung von Boc-5-aminopentanonsäure (1,09 mMol) und NMM (3,27 mMol) in trockenem DMF (6 ml) wurde HATU (1,14 mMol) hinzugefügt und das Reaktionsgemisch bei Raumtemperatur für 0,5 Stunden gerührt. Eine Lösung von Verbindung 3 (1,20 mMol) in trockenem DMF (1 ml) wurde hinzugefügt und das Reaktionsgemisch bei Raumtemperatur für 18 Stunden gerührt. Die Reaktion wurde mit EtOAc (50 ml) verdünnt und mit 1,0 M HCl (2 × 50 ml) und D.I. H<sub>2</sub>O (2 × 50 ml) gewaschen. Die organischen Extrakte wurden getrocknet (MgSO<sub>4</sub>) und zur Trocknung abgedampft. Der Rest wurde Säulenchromatographie unterzogen, um 420 mg (91%) von Verbindung 4 zu ergeben: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8,65 (d, 1H), 7,88 (d, 1H), 7,65 (m, 2H), 7,48 (t, 1H), 6,73 (br s, 1H), 5,85 (m, 1H), 5,55 (m, 1H), 5,23 (q, 2H), 4,55 (d, 2H), 2,80 (m, 2H), 2,05 (t, 2H), 1,33 (s, 9H), 1,20–1,30 (m, 4H).

[0621] C. Eine Lösung von Verbindung 4 (0,9 mMol) in HCl-1,4-Dioxan (20 mMol) wurde bei Raumtemperatur

für 2 Stunden gerührt. Das Reaktionsgemisch wurde konzentriert, in MeOH und Toluol gelöst und nochmals konzentriert (5 × 5 ml) um 398 mg (quantitativ) der Verbindung 5 zu ergeben: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8,75 (d, 1H), 7,88 (d, 1H), 7,65 (m, 2H), 7,51 (t, 1H), 7,22 (m, 2H), 5,85 (m, 1H), 5,57 (m, 1H), 5,23 (q, 2H), 4,55 (d, 2H), 2,80 (m, 2H), 2,71 (m, 2H), 2,07 (s, 2H), 1,40–1,48 (br s, 4H).

[0622] D. Zu einer Lösung von Verbindung 21 (0,48 mMol, hergestellt gemäß Beispiel 20) und NMM (1,44 mMol) in trockenem DMF (3 ml) wurde HATU (0,50 mMol) hinzugefügt und das Reaktionsgemisch bei Raumtemperatur für 0,5 Stunden gerührt. Eine Lösung von Verbindung 5 (0,51 mMol) in trockenem DMF (3 ml) wurde hinzugefügt und die Reaktion bei Raumtemperatur für 18 Stunden gerührt. Das Reaktionsgemisch wurde mit EtOAc (75 ml) verdünnt und mit 5% Na<sub>2</sub>CO<sub>3</sub> (3 × 50 ml) gewaschen. Die organischen Extrakte wurden getrocknet (MgSO<sub>4</sub>) und das Lösungsmittel verdampft, um 281 mg (78%) von Verbindung 6 zu ergeben: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8,65 (d, 1H), 8,17 (d, 1H), 7,82–7,95 (m, 4H), 7,68 (m, 3H), 7,50 (t, 1H), 6,92 (d, 1H), 5,85 (m, 1H), 5,57 (m, 1H), 5,20 (q, 2H), 4,55 (d, 2H), 4,30 (q, 1H), 4,05 (q, 2H), 2,95 (m, 4H), 2,80 (m, 2H), 2,72 (m, 2H), 2,05 (s, 3H), 2,01 (t, 2H), 1,58–1,77 (m, 3H), 1,50 (m, 4H), 1,30 (q, 3H), 1,17–1,40 (m, 9H).

[0623] E. Zu einem Gemisch von Verbindung 6 (0,36 mMol) in THF (4 ml), wurden 1 M NaOH (1 mMol) hinzugefügt und die Reaktion bei Raumtemperatur für 2 Stunden gerührt. Das Reaktionsgemisch wurde auf pH 7,0 mit 1,0 M HCl (1 ml) angesäuert und das Lösungsmittel abgedampft, um Verbindung 7 zu ergeben (quantitativ): <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8,65 (d, 1H), 8,17 (d, 1H), 7,82–7,95 (m, 4H), 7,68 (m, 3H), 7,50 (t, 1H), 6,92 (d, 1H), 5,52 (m, 1H), 4,30 (q, 1H), 4,05 (q, 2H), 2,95 (m, 4H), 2,80 (m, 2H), 2,72 (m, 2H), 2,05 (s, 3H), 2,01 (t, 2H), 1,58–1,77 (m, 3H), 1,50 (m, 4H), 1,30 (q, 3H), 1,17–1,40 (m, 9H).

[0624] F. Zu einer Lösung von Verbindung 7 (0,04 mMol) und NMM (0,12 mMol) in trockenem DMF (0,4 ml), wurde HATU (0,044 mMol) hinzugefügt und die Reaktion bei Raumtemperatur für 0,5 Stunden gerührt. Allylamin (0,12 mMol) wurde hinzugefügt und das Reaktionsgemisch bei Raumtemperatur für 5 Stunden gerührt. Das Reaktionsgemisch wurde mit EtOAc (15 ml) verdünnt und mit 5% Na<sub>2</sub>CO<sub>3</sub> (3 × 10 ml) gewaschen. Die organischen Extrakte wurden getrocknet (MgSO<sub>4</sub>) und das Lösungsmittel abgedampft, um 15 mg (49%) von Verbindung 8 zu ergeben: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8,49 (d, 1H), 8,17 (d, 1H), 7,82–7,95 (m, 4H), 7,68 (m, 3H), 7,50 (t, 1H), 6,92 (d, 1H), 5,72 (m, 1H), 5,50 (m, 1H), 5,03 (q, 2H), 4,37 (d, 2H), 4,30 (q, 1H), 4,05 (q, 2H), 2,95 (m, 4H), 2,80 (m, 2H), 2,72 (m, 2H), 2,05 (s, 3H), 2,01 (t, 2H), 1,581,77 (m, 3H), 1,50 (m, 4H), 1,30 (q, 3H), 1,17–1,40 (m, 9H).

#### BEISPIEL 19

[0625] Die Sequenz von Reaktionen A → G wie in diesem Beispiel 19 beschrieben, wird in **Fig. 20A** und **20B** dargestellt. Die in diesem Beispiel angegebenen Verbindungsnummern beziehen sich auf die Verbindungen der selben Nummern, wie in **Fig. 20A** und **20B**.

[0626] A. Zu einer Lösung von Fmoc-Lys(Boc)-OH (Verbindung 9, 33,8 mMol) in CHCl<sub>3</sub> (200 ml), wurde Diisopropylethylamin (67,5 mMol) und Allylbromid (67,5 mMol) hinzugefügt. Das Reaktionsgemisch wurde für 6 Stunden refluxiert, bei Raumtemperatur für 16 Stunden gerührt, verdünnt mit CHCl<sub>3</sub>, gewaschen mit 1,0 M HCl (2 × 150 ml), gesättigtem NaHCO<sub>3</sub> (1 × 150 ml) und D.I. H<sub>2</sub>O (2 × 150 ml). Die organischen Extrakte wurden getrocknet (MgSO<sub>4</sub>) und das Lösungsmittel abgedampft, um Verbindung 10 zu ergeben.

[0627] Zu einer Lösung von Verbindung 10 in CHCl<sub>3</sub> (90 ml) wurde Pyrrolidin (10 äq.) hinzugefügt und die Reaktion wurde bei Raumtemperatur für 2,5 Stunden gerührt. Das Reaktionsgemisch wurde mit CHCl<sub>3</sub> (150 ml) verdünnt und mit gesättigtem NaHCO<sub>3</sub> (3 × 250 ml) gewaschen. Die organischen Extrakte wurden getrocknet (MgSO<sub>4</sub>) und das Lösungsmittel abgedampft. Der Rest wurde Säulenchromatographie unterzogen (EtOAc/McOH) um 6,52 g (67%) von Verbindung 11 zu ergeben: <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 5,90 (m, 1H), 5,27 (m, 2H), 4,60 (d, 2H), 3,48 (t, 1H), 3,10 (d, 2H), 1,40–1,78 (m, 9H), 1,40 (s, 9H).

[0628] B. Zu einer Lösung von N-Methylisonipecotinsäure (1,60 mMol) und N-Methylmorpholin (4,80 mMol) in trockenem DMF (5 ml) wurde HATU (1,67 mMol) hinzugefügt. Nach 0,5 Stunden wurde eine Lösung der Verbindung 11 (1,75 mMol) in trockenem DMF (2 ml) hinzugefügt und das Reaktionsgemisch bei Raumtemperatur für 18 Stunden gerührt. Das Reaktionsgemisch wurde mit CH<sub>2</sub>Cl<sub>2</sub> (60 ml) verdünnt und mit gesättigtem Na<sub>2</sub>CO<sub>3</sub> (3 × 40 ml) gewaschen. Die organischen Extrakte wurden getrocknet (MgSO<sub>4</sub>) und das Lösungsmittel abgedampft. Der Rest wurde Säulenchromatographie unterzogen (CH<sub>2</sub>Cl<sub>2</sub>/McOH/ Triethylamin), um 580 mg (88%) von Verbindung 12 zu ergeben: <sup>1</sup>H NMR (DMSO): δ 8,12 (d, 1H), 6,77 (t, 1H), 5,90 (m, 1H), 5,27 (m, 2H), 4,53 (d, 2H), 4,18 (m, 1H), 2,62–2,90 (m, 5H), 2,13 (s, 3H), 1,85 (m, 2H), 1,57 (m, 5H), 1,35 (s, 9H), 1,00 (t, 2H).

[0629] C. Ein Gemisch von Verbindung 12 (1,39 mMol) in HCl-1,4-Dioxan (20 mMol) wurde bei Raumtemperatur für 4 Stunden gerührt. Das Reaktionsgemisch wurde konzentriert, in McOH gelöst, mit Toluol co-abgedampft (5 × 5 ml), um 527 mg (quantitativ) von Verbindung 13 zu ergeben: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8,12 (d, 1H), 6,77 (t, 1H), 5,90 (m, 1H), 5,27 (m, 2H), 4,53 (d, 2H), 4,18 (m, 1H), 2,65–3,00 (m, 8H), 2,23 (s, 3H), 1,85 (m, 2H), 1,57 (m, 5H), 1,00 (t, 2H).

[0630] D. Zu einer Lösung von 4-Ethoxybenzoesäure (1 äq.) in trockenem DMF wird NMM (3 äq.) und HATU

(1,05 äq.) hinzugefügt. Nach 0,5 Stunden wird eine Lösung von Verbindung 13 in trockenem DMF hinzugefügt. Nach der Vervollständigung der Reaktion und der basischen Aufarbeitung wird die Verbindung 14 isoliert und gereinigt.

[0631] E. Zu einer Lösung von Verbindung 14 in THF wird 1N NaOH hinzugefügt und das Reaktionsgemisch bei Raumtemperatur gerührt. Nach der Vervollständigung der Reaktion und Ansäuerung wird die Verbindung 15 isoliert.

[0632] F. Zu einer Lösung von Verbindung 15 (1 äq.) in trockenem DMF, wird NMM (3 äq.) und HATU (1,05 äq.) hinzugefügt. Nach 0,5 Stunden wird eine Lösung von Verbindung 21 (ANP-Allylester, gemäß Beispiel 20 hergestellt) in trockenem DMF hinzugefügt. Nach der Vervollständigung der Reaktion und der basischen Aufarbeitung wird die Titelverbindung 16 isoliert und gereinigt.

[0633] G. Zu einer Lösung von Verbindung 16 in THF wird 1N NaOH hinzugefügt und das Reaktionsgemisch bei Raumtemperatur gerührt. Nach der Vervollständigung der Reaktion und der Ansäuerung wird die Verbindung 17 isoliert.

#### BEISPIEL 20

[0634] Die Sequenz der Reaktion A bis D, wie in diesem Beispiel 20 beschrieben, wird in **Fig. 21** dargestellt. Die wie in diesem Beispiel angegebenen Verbindungsnummern, sowie in den Beispielen 18 und 19 beziehen sich auf die Verbindungen mit derselben Nummer wie in **Fig. 21**.

[0635] A. Zu einer Lösung von 4-Ethoxybenzoesäure (7,82 mMol) und N-Methylmorpholin (20,4 mMol) in  $\text{CH}_2\text{Cl}_2$  (10 ml) wurde HATU (7,14 mMol) hinzugefügt. Nach 0,25 Stunden wurde eine Lösung der Verbindung 11 (6,8 mMol) in  $\text{CH}_2\text{Cl}_2$  (6 ml) hinzugefügt und das Reaktionsgemisch bei Raumtemperatur für 18 Stunden gerührt. Die Reaktion wurde mit  $\text{CH}_2\text{Cl}_2$  (150 ml) verdünnt und mit 1,0 M HCl (3 × 50 ml) und gesättigtem  $\text{NaHCO}_3$  (3 × 50 ml) gewaschen. Die organischen Extrakte wurden getrocknet ( $\text{MgSO}_4$ ) und das Lösungsmittel abgedampft. Der Rest wurde Säulenchromatographie unterzogen ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$ ), um 2,42 g (82%) von Verbindung 18 zu ergeben:  $^1\text{H NMR}$  ( $\text{CDCl}_3$ ):  $\delta$  7,78 (d, 2H), 6,91 (d, 2H), 6,88 (d, 1 H), 5,83–5,98 (m, 1 H), 5,21–5,38 (m, 2H), 4,80 (q, 1 H), 4,66 (d, 2H), 4,06 (q, 2H), 3,11 (q, 2H), 1,90–2,04 (m, 1H), 1,68–1,87 (m, 1H), 1,39 (t, 3H), 1,34 (s, 9H), 1,32–1,58 (m, 4H).

[0636] B. Ein Gemisch von Verbindung 18 (5,5 mMol) in HCl-1,4-Dioxan (14,3 mMol) wurde bei Raumtemperatur für 1 Stunde gerührt. Das Reaktionsgemisch wurde konzentriert, in MeOH gelöst, azetropiert mit Toluol und wiederum konzentriert (5 × 5 ml), um eine quantitative Ausbeute von Verbindung 19 zu ergeben.

[0637] C. Zu einer Lösung von N-Methylisonipekotsäure (6,21 mMol) in trockenem DMF (15 ml) wurde NMM (21,6 mMol) und HATU (5,67 mMol) hinzugefügt. Nach 0,5 Stunden wurde eine Lösung von Verbindung 19 (5,4 mMol) in trockenem DMF (10 ml) hinzugefügt und die Reaktion bei Raumtemperatur für 18 Stunden gerührt. Das Reaktionsgemisch wurde auf pH 12 mit 1N NaOH (20 ml) gebracht und mit  $\text{CHCl}_3$  (2 × 200 ml) extrahiert. Die organischen Extrakte wurden getrocknet ( $\text{MgSO}_4$ ) und das Lösungsmittel abgedampft, um 2,2 g (89%) von Verbindung 20 zu ergeben:  $^1\text{H NMR}$  ( $\text{DMSO}-d_6$ ):  $\delta$  8,52 (d, 1H), 7,84 (d, 2H), 7,72 (t, 1H), 6,95 (d, 2H), 5,80–5,95 (m, 1H), 5,18–5,31 (dd, 2H), 4,58 (d, 2H), 4,37 (q, 1H), 4,08 (q, 2H), 3,01 (d, 2H), 2,08 (s, 3H), 1,95 (m, 1H), 1,63–1,82 (m, 4H), 1,51 (m, 4H), 1,32 (t, 3H), 1,22–1,41 (m, 6H).

[0638] D. Zu einer Lösung von Verbindung 20 (4,4 mMol) in THF (10 ml), wird 1N NaOH (4,4 mMol) hinzugefügt und das Reaktionsgemisch bei Raumtemperatur für 1 Stunde gerührt. Die Reaktion wurde konzentriert, in THF/Toluol (2 × 5 ml) gelöst, konzentriert, in  $\text{CH}_2\text{Cl}_2$ /Toluol (1 × 5 ml) gelöst und noch mal konzentriert, um eine quantitative Ausbeute von Verbindung 21 zu ergeben:  $^1\text{H NMR}$  ( $\text{DMSO}-d_6$ ):  $\delta$  7,76 (d, 2H), 6,96 (d, 2H), 4,04 (q, 2H), 3,97 (d, 1H), 2,97 (d, 2H), 2,64 (d, 2H), 2,08 (s, 3H), 1,95 (m, 1H), 1,58–1,79 (m, 4H), 1,44 (m, 6H), 1,30 (t, 3H), 1,11–1,35 (m, 4H).

#### BEISPIEL 21

[0639] Die Synthese der CMSTs (spaltbarer Massenspektrometrie-nachweisbarer Marker) können, wie in **Fig. 22** beschrieben, auf einem kombinatorischen Ansatz basiert sein. Der allgemeine Ansatz ist so aufgebaut, um mit den Entwicklungen der Massenspektrometer und den Veränderungen und Verbesserungen in den Ionisationstechnologien kompatibel zu sein. Ein zentrales Gerüst wird zuerst auf seine Kompatibilität mit dem Typ von Ionisation getestet, der in dem Verfahren verwendet werden wird. Es ist wichtig, daß das Gerüst nicht gegenüber Fragmentierung, Hitzedegradation, oder Dimerisierung zur Adduktbildung empfindlich ist. Mit den augenblicklichen APCUQuadrupol-Massenspektrometern werden ungefähr 400 Markern das Spektrum füllen, wobei isotopische Kontamination berücksichtigt wird, die zu einer minimalen Beabstandung von ungefähr 4 AMU zwingt.

[0640] Die CMSTs wurden auf eine modulare Weise aufgebaut, so daß die Macker mit einem kombinatorischen Chemie-Ansatz aufgebaut werden können. Es gibt 5 "Module" (funktionell getrennte Atomgruppen) in dem vollständigen Marker-Oligonukleotid. Das erste Modul ist das Oligonukleotid (ODN), das von jeder Länge

und Sequenz sein kann und das ein 5'-Hexylamin zur Kopplung an den Linker besitzt, der bevorzugterweise ein photospaltbarer Linker ist. Das ODN kann entweder als eine Sonde für die Hybridisierung oder als ein Primer in einem Verlängerungs-, Ligationss- oder Enzym-basierten Assay sein. Das zweite Modul ist der Linker, bevorzugterweise ein photolabiler Linker, der den Marker an die Oligonukleotidsonde oder den Primer bindet. In dem hier genau beschriebenen Ansatz ist der photolabile Linker ein o-Nitrobenzylaminosäurederivat (siehe Lebl M. et al. "Dynamic database of references in molecular diversity", der die Internet-Adresse <http://www.5z.com> für eine Zusammenfassung hat). Das dritte Modul ist der Ionisationsverstärker. Dieses Modul ist das Gerüst, auf das der CMST synthetisiert wird und stellt im allgemeinen Funktionalität zur Verfügung, die für den Typ von Ionisationsverfahren, das verwendet wird (d. h. APCI (positiv oder negativer Modus) Elektrospray, MALDI, usw.), spezifisch ist. Das vierte Modul ist der Gesamtmasseinsteller, womit die Masse in großen Schritten von 200–500 amu verändert werden kann. Dieses Modul erlaubt die Wiederverwendung der variablen Gewichtseinsteller. Das fünfte Modul ist der variable Gewichtseinsteller, auch als variable Masseneinheiten (VMU) bezeichnet. Die variablen Gewichtseinsteller sind chemische Seitenketten, die zu dem Markergerüst hinzugefügt werden. Diese variablen Gewichtseinsteller stellen das Gewicht der CMSTs fein ein. Das Gewicht der Marker ist mindestens jede 4 amu voneinander beabstandet, um überlappende Spektren aufgrund der isotopischen Kontamination zu vermeiden. Dieselben VMU-Seitenarme können wiederholt mit den verschiedenen Typen von Ionisationsgerüsten verwendet werden. Zusammengefaßt sind das Ionisationsmodul und die variablen Masseinsteller so aufgebaut, um ein vorhersagbares Verhalten in dem MSD zu bewirken, die Photospaltung des photolabilen Linkers ist schnell (unten beschrieben) und die CMST/Oligonukleotidkonjugate sind mit PCR und HPLC und anderen in Assay-Formaten gefundenen Manipulationen kompatibel.

[0641] Eine genaue synthetische Route der Marker ist unten und in **Fig. 22** beschrieben. Die synthetische Route in Richtung eines CMST beginnt mit der Esterifizierung des photosensitiven ANP-Linkers (3-Amino-3-(2-nitrophenyl)propionsäure, (1) um das Ethylesterhydrochlorid (2) mit 84% Ausbeute zu ergeben. Ein wichtiger Schritt in dem Verfahren ist die enzymatische Transformation (2), um den Ethylester als ein einzelnes Isomer zur Verfügung zu stellen. Nachdem das Ethylesterhydrochlorid zum freien Amin basifiziert wurde und konzentriert wurde, wurde der ölige Rest in pH 7 Phosphatpuffer aufgenommen und auf neutralen pH mit 2N HCl eingestellt. Das Amano PS Esteraseenzym wurde als eine Phosphatpufferschlämme hinzugefügt. Nach der Vervollständigung der Reaktion entfernte eine basische Aufarbeitung das hydrolysierte ANP-Nebenprodukt (4) und der Einzel-Isomerethylester (3 > 99% e.e.) wurde zurückerhalten (92% des erhältlichen Materials).

[0642] Das Koppeln von (3) mit einem α-BOC-e-alloc-lysin (5), unter der Verwendung von EDAC und HOBT, ergab das geschützte ANP-Lysin (6) in 91% Ausbeute. Das Entfernen von BOC mit TFA stellt den Amino-e-alloc-lysin ANP-Ester (7) als einen weißen Feststoff zur Verfügung. Methyl-isonipecotische Säure Hydrochloride wurde an 7 gekoppelt, unter der Verwendung von EDAC und Triethylamin, um ungereinigte Alloc-geschützte Kernstruktur (8) zu ergeben, die mit Diethylamin, Triphenylphosphin und Palladiumacetat bei 50°C geschützt wurde. Die sich ergebende Kernstruktur (9) wurde aus dem Reaktionsgemisch kristallisiert und durch Filtration in 95% Ausbeute als ein gelber Feststoff zurückgehalten.

[0643] Eine Vielzahl von Carboxylsäuren, die als variable Masseneinheiten (VMU) bezeichnet wurden, wurden an (9) gekoppelt, unter der Verwendung von HATU und N-Methylmorpholin. Ein Set von VMU's wurden so aufgebaut, um Gleich-Massenmarker zur Verfügung zu stellen und 4 amu Beabstandung, um eine isotopische Kontamination zu vermeiden. Die Folgenden wurden als Ausschlußkriterien bei der Auswahl bestimmter VMU's der Zielmassen verwendet: 1) Funktionalitäten, die mit der synthetischen Sequenz inkompatibel waren (z. B. Ester); 2) Elemente mit multiplen Isotopen (z. B. Cl, Br, S); 3) Funktionalitäten, die vielleicht zu störenden Photoprozessen führen (Jodid, Acyl- und Aryl-phenone); 4) racemische Säuren und 5) die Erhältlichkeit von Lieferanten.

[0644] Nach der Reinigung durch Säulenchromatographie wurde der CMST-Ethylester (10) in veränderlichen Ausbeuten zurückgehalten. Basische Hydrolyse von (10) mit NaOH ergab die CMST-Säure (11) in quantitativer Ausbeute. Der letzte Schritt, die Bildung des aktivierten Esters, wurde unter Verwendung von Tetrafluorphenoltrifluoroacetat und Hunnig's Base erreicht und führte zu dem CMST-TFP-Ester (12) in verschiedenen Ausbeuten.

[0645] Es ist bequem, die Marker an das 5'-Ende der Oligonukleotide anzubringen (im allgemeinen), so daß das 3'-Hydroxyl in der Polymerasekettenreaktion verlängert werden kann oder für andere enzymatische Modifikationen zugänglich ist. Auch, wenn direkt als Sonden wie in der Anwendung hier beschrieben, verwendet, werden die Marker bevorzugterweise an das 5'-Ende der Oligonukleotide angebracht. Die durch Lukhtanov et. al., beschriebenen Richtlinien "Oligodeoxyribonucleotides with conjugated dihydropyrroloindole oligopeptides: preparation and hybridization properties", *Bioconjug Chem* 6(4): 418–26, Jul–Aug, 1995, können befolgt werden, um die markierten Oligonukleotide aus CMSTs und Oligonukleotiden herzustellen.

## BEISPIEL 22

[0646] Eine PHRED-Photospaltungseinheit wird zwischen die HPLC und das Massenspektrometer oder ei-



nen Autoinjektor und den MSD plziert. "PHRED", das für Photochemical Reactor for Enhanced Detection steht, und erhältlich ist von Aura Industries, Staten Island, NY (erhältlich mit sowohl einer 254 nm und 366 nm Lampe, die 254 nm Lampe wurde verwendet). Eine in-line-Vorrichtung wird bevorzugterweise zwischen die Abtrenninstrumentierung (z. B. HPLC oder Gel) und der Detektor plziert. Die Schnittstelle weist bevorzugterweise die folgenden Eigenschaften auf: Die Fähigkeit, die DNA-Fragmente zu diskreten Zeitintervallen zu sammeln, die DNA-Fragmente zu konzentrieren, die DNA-Fragmente aus den Elektrophoresepuffern und dem Milieu zu entfernen, die MW-Identifizierer von dem DNA-Fragment abzuspalten, die MW-Identifizierer von dem DNA-Fragment abzutrennen, das DNA-Fragment zu verwerfen, den Marker in eine flüchtige Lösung zu plazieren, den Marker zu verflüchtigen und ionisieren, was den Marker in das Massenspektrometer einführt.

[0647] Eine geeignete Konfiguration der Photospaltungsvorrichtung ist eine 300 cm lange, 8 Watt, UV-keimabtötende Lampe (G8T50) (mit einer Emission um 366 nm) unter die eine 801 Spule von 0,01 Inch ID Tefzel-Röhre plziert wird. Flußraten von 800 l pro Minute sind im allgemeinen geeignet. Die Lösungszusammensetzungen, die mit APCUMS kompatibel sind und die nur geringe Konzentrationen von Acetonitril und Puffer, wie zum Beispiel Tris-HCl enthaltend, sind bevorzugt. Mit diesem Aufbau gibt es kein Erfordernis, die DNA von dem Marker vor dem Ionisationsschritt abzutrennen. Der photolabile Linker spaltet sich sehr schnell unter diesen Bedingungen. Die Wärmequelle in der APCI-Kammer trägt zu der Spaltung des photolabilen Linkers bei.

[0648] Durch Variieren der Länge der Tefzel-Röhrenspule unter der UV-Quelle und Konstanthalten der Flußrate wurde die Verweildauer unter der UV-Quelle von den 0,75 bis 6 Sekunden variiert. Der Antwortfaktor (der integrierte Ionenstrom, der pro Mol Marker produziert wurde, das in den Flußstrom injiziert wurde, in der "Flußinjektionsanalyse" (FIA)) der Marker wurde für einen Pool von 6 Markierungen bestimmt, die alle an einer einzelnen Oligonukleotidsequenz angebracht waren (21-mer, M13 Vorwärtssequenzierungsprimer). Der Antwortfaktor ist das Integral der Effizienz der Ionisation, Ioneneinführung in die Vakuumkammer und dann anschließender Nachweis durch den MSD. Jede Markierung in dem Pool war bei einer Konzentration von 100 fmol pro µl vorhanden. Das Verdünnungsmittel war tRNA (Boehringer Mannheim) bei einer Konzentration von 1 µg/ml in HPLC-Grad Wasser. In Tabelle 3 sind der Antwortfaktor und die Prozentsätze des beobachteten Signals relativ zu dem längsten Aussetzen (6 Sekunden) auf einer Pro-Markierungsbasis aufgelistet. Die Ergebnisse zeigen an, daß der Marker schnell wird (weniger als 2 Sekunden) von dem befestigten Oligonukleotid abgespalten wird. Im Zeitrahmen von 1,9 bis 6 Sekunden gab es wenig Abnahme im beobachteten Antwortfaktor mit 6 verschiedenen markierten ODNs. Bei der kürzesten Aussetzungszeit, die getestet wurde (0,75 Sekunden) gab es bis zu einer 25%igen Abnahme in dem beobachteten Antwortfaktor.

TABELLE 3

Marker MW	10 µl	25 µl	50 µl	80 µl
447	4300 81%	4400 83%	5300 100%	5300 100%
455	4900 92%	5200 98%	6400 120%	5300 100%
479	3600 75%	4100 85%	4900 102%	4800 100%
503	4600 74%	5700 92%	6700 108%	6200 100%
507	5200 85%	5400 88%	6500 104%	6100 100%
511	4100 78%	5100 98%	5500 105%	5200 100%

[0649] Es gab eine ungefähr 40%-ige Abnahme in dem Antwortfaktor der sechs Marker, wenn die Lampe in der Photospaltungseinheit (Daten nicht gezeigt). Es wurde nachgewiesen, daß der photolabile Linker thermolabil ist und der Marker offensichtlich während des APC-Ionisationsschrittes abgespalten wird, wenn die Verdampfungstemperatur 450°C ist.

## BEISPIEL 23

## KOLLEKTIVES MARKERVERHALTEN UND RELATIVE STABILITÄT

[0650] Die Marker wurden so aufgebaut, um ein einzelnes parenterales Ion zur Verfügung zu stellen, das nicht in Tochter-Ionen zerfiel oder Addukte bildete. Die Antwort der Marker wurde für einen Pool von 43 Markern bestimmt, die alle an eine einzelne Oligonukleotidsequenz befestigt waren (20-mer, M13 Sequenz). Jeder Marker in dem Pool war bei einer Konzentration von 100 fmol pro µl vorhanden (die Konzentration wurde bestimmt

durch Anwenden eines Verdünnungsfaktors auf eine Oligo/Makervorratslösung, die spektrophotometrisch bei 260 nm gemessen wurde. Eine Korrektur auf den Markerbeitrag zur Absorption bei 260 nm wurde nicht verwendet. Das Verdünnungsmittel war tRNA (Boehringer Mannheim) bei einer Konzentration von 1 µg/ml in HPLC-Grad Wasser. Der Pool wurde bei 4°C im Dunkeln gelagert. Die CMST-markierten Oligonukleotide können unter normalen Labor-Lichtbedingungen gehandhabt werden, ohne das Auftreten von wahrnehmbarer Degradation.

[0651] Vor der Analyse wurde ein 55 µl Aliquot aus der Vorratslösung entfernt und in ein 200 µl Polypropylen-Autosamplergefäß plaziert und durch Falten geschlossen. Fünf Injektionen von jeweils 5 µl aus dem Pool und 3 Injektionen von jeweils 5 µl von dem Verdünnungsmittel wurden durch den HPLC HP1100 ALS durchgeführt. Die APCI-MS-Kammer Parameter waren wie folgt: 20 PSI für den Verneblungsdruck, die Verdampfungstemperatur war 450°C, der Trocknungsgasfluß war 3 l/min, die Trockengastemperatur war 350°C, Coronastrom war 4 uA, Fragmentorspannung 125 V, die Verstärkung wurde auf 1 gesetzt und die Peakweite 0,07 Minuten. Die Flußrate war 0,8 ml/min, das "Totvolumen" der Photospaltungseinheit war 80 µl (0,01" ID Tefzel). Die Lampe der Photospaltungseinheit arbeitete bei 366 nm. Jeder Marker wurde durch die Extraktion des SIM-Ions aus dem TIC quantifiziert. Die Peaks wurden unter den folgenden Parametern integriert: Steigungssensitivität bei 2500, Minimum Peakfläche bei 800, Minimum Peakhöhe bei 100, Peakbreite bei 0,15, die Schultereinstellung bei "aus". Die Peakfläche für jeden Marker wurde für alle fünf Injektionen innerhalb eines Experiments aufgezeichnet und die durchschnittliche Fläche wurde zusammen mit der Standardabweichung und dem Variationskoeffizienten berechnet. Durchschnittliche Bereiche, Standardabweichungen und CVs wurden auch für einen einzelnen Tag, über eine zwei Tage Zeitdauer und für eine drei Tage-Zeitdauer hinweg berechnet. Innerhalb eines einzelnen Tages variierte der Koeffizient der Variation zwischen 2,0% und 9,9% zwischen verschiedenen Markern. Über eine drei Tage-Zeitdauer variierte der Koeffizient der Variation zwischen 4,0 und 9,8% zwischen verschiedener, Markern. Die Marker sind daher im Hinblick auf Lagerung und Poolen stabil, wie in Tabellen 4a, 4b und 4c gezeigt.

TABELLE 4a  
1-Tag

Tag MW	CV @ 10 fmol/Injektion	100 fmol/Injektion
447	2,0	3,7
455	4,0	6,5
479	6,9	7,9
503	9,9	2,9
507	4,1	1,8
511	4,1	4,0

TABELLE 4b  
2-Tage

Tag MW	CV @ 10 fmol/Injektion	100 fmol/Injektion
447	3,0	4,1
455	3,8	8,6
479	4,8	6,7
503	1,2	5,6
507	5,5	8,7
511	4,5	5,6

TABELLE 4e  
3-Tage

Tag MW	CV @ 10 fmol/Injektion	100 fmol/Injektion
447	3,3	4,9
455	4,0	3,9
479	5,6	6,7
503	9,8	8,9
507	2,0	4,8
511	3,8	6,7

## Sensitivität und untere Nachweisgrenze

[0652] Spannungsmassen-produzierte Quadrapole weisen Sensitivitäten auf, die mit Fluoreszenz-basierter Sequenzierung vergleichbar sind. Die Sensitivität kann im Hinblick auf die untere Nachweisgrenze der Masse des CMST angegeben werden. Wir definieren die untere Nachweisgrenze als 3 Standardabweichungen oberhalb des Hintergrunds des Assay-Systems. Die untere Nachweisgrenze des Pools von 43 Konjugaten wurde für den 30-Ion SIM Modus bestimmt.

[0653] Ein Set von 10 zweifachen Verdünnungen wurde in dem tRNA/Wasser-Verdünnungsmittel präpariert, um 500, 250, 125, 62,5, 31,2, 15,6, 7,8, 3,9, 1,9 und 0,9 Femtomol von Material pro Injektion (5–10 µl) zu ergeben. Die Daten wurden in dem 30-Ion SIM Modus unter oben beschriebenen Bedingungen erhalten: Die LLD für jeden Marker ist in Tabelle 5 gezeigt.

TABELLE 5  
UNTERE NACHWEISGRENZE FÜR 43 CMSTs.

Marker MW	LLD	Signal bei LLD
367	30	1970
371	30	1707
375	15	1318
379	15	1642
383	30	2585
387	15	1301
391	15	1554
395	15	1960
403	8	1784
407	15	2266
411	15	2285

415	8	1362
419	8	1260
423	8	1530
427	15	1607
431	8	1606
439	8	2797
443	4	1576
447	4	1144
459	8	1881
463	8	1179
467	8	1625
471	2	1094
475	4	1664
479	2	2626
483	8	1528
487	8	1556
495	15	1894
499	15	1465
503	4	1208
507	8	1843
511	8	1453
515	8	1276
519	8	1042
523	15	1690
527	8	815
531	8	1521
535	8	1276
539	4	1651
543	15	1076

[0654] Die LLD für dieses bestimmte Set von Markern ist ungefähr 4–30 Femtomol pro Injektion im Anschluß an den Photospaltungsschritt des Oligonukleotids: ( $30 \times 10^{-15}$  Mole @ 500 MW  $\rightarrow 5 \times 10^{-11}$  Gramm  $\rightarrow$  50 Picogramm Marker,  $100 \times 10^{-15}$  Mole bei  $330 \times 400$  nt  $\times 2 = 3 \times 10^{-8}$  Gramm = 30 Nanogramm eines 400 nt doppelsträngigen PCR-Produkts). Daher, unter der Annahme, daß eine durchschnittliche 25 µl PCR-Reaktion 300 ng eines doppelsträngigen PCR-Produkts enthält, wird ungefähr 1/10 dieses Produkts dazu verwendet, um ein Signal zu erzeugen, das statistisch oberhalb des Hintergrunds der Messung liegt. Zu Vergleichszwecken wurden ungefähr 100–200 ng DNA pro Spur auf dem ABI 377 Sequenzierer verwendet.

[0655] Die untere Nachweisgrenze des Markierungssystems beeinträchtigt den Grad an Multiplexierung, das in einer einzelnen Injektion erreicht werden kann. Im Augenblick haben wir kein Problem mit dem Meßprozeß festgestellt, wenn bis zu 10 µg an PCR-Produkt pro Injektion verwendet wird. Die Implikation dieser Zahl ist, daß 10.000/30 oder ungefähr 300 Reaktionen unter der Verwendung des Markierungssystems der vorliegenden Erfindung multiplexiert werden können. Daher ist für das HP-MSD 1.100 das Maß von Multiplexierung ungefähr dasselbe, wie die Marker, die in das Spektrum des Quadrupol MS plaziert werden können.

#### Markerinterferenz

[0656] Es gibt wenig oder keine Interferenz im Hinblick auf die Antwort (wobei Antwort der meßbare Ionenstrom ist, der das Ausmaß an Ionisation, Ausmaß von Fragmentierung oder Abbau, Adduktbildung, usw. widerspiegelt) bei der Multiplexierung großer Zahlen von CMSTs der vorliegenden Erfindung. Die Multiplexierung von Markern beeinträchtigt nicht die Antwort von einzelnen Markern und daher beeinträchtigt die Multiplexierung nicht die Ionisation oder den gesamten Ionenstrom.

#### Antwortfaktor als eine Funktion des Injektionsvolumens

[0657] Wie oben beschrieben, wurde ein Pool von 43 Konjugaten gemessen, um den Antwortfaktor als eine Funktion des Injektionsvolumens zu bestimmen. 50 fmol pro Injektion und 500 fmol pro Injektion wurden in 5, 10, 20, 50 und 100 µl Volumen gemessen. Die APCI-MS Parameter waren identisch zu denjenigen, die in dem Markerstabilitätsexperiment verwendet wurden. Die in Tabelle 4 gezeigten Werte sind ein Mittelwert von 5-Replikatinjektionen für jedes gemessene Volumen. Es gab keinen Anstieg im RF zwischen 5 µl und 10 µl Volumina. Bei den 20 µl Injektionsvolumina war der RF 90–97% des 5 µl Volumens. Wenn 50 µl Volumina gemessen wurden, war der RF 54%–75% des 5 µl Volumens. Es gab kein nachweisbares Signal unter der Verwendung der 100 µl Injektionsvolumina.

#### BEISPIEL 24

##### P4502D6 POLYMORPHISMUSNACHWEIS

[0658] Die hier beschriebene CMST-Technologie-Plattform kann dazu verwendet werden, den CYP2D6 Polymorphismus zu messen, der für den Stoffwechsel von Debrisoquin-4-hydroxylase verantwortlich ist. Dieses bestimmte P450 Cytochrom ist im Stoffwechsel von mehr als 30 Wirkstoffen und xenobiotischen Verbindungen wichtig. CYP2D6 (P450-2D6) ist geschätzt, wie für den Stoffwechsel von ungefähr 25% der im Augenblick verschriebenen Wirkstoffe verantwortlich. Die CYP2D6 ist auch als Debrisoquin/Spartein-Hydroxylase bekannt. Sache, Am. J. Human Genetics, 60: 284–295, 1997, hat geschätzt, daß bis zu 10% aller Kaukasier schlechte Metabolisierer sind, aufgrund entweder inaktivierenden Mutationen in beiden Allelen oder dem vollständigen Fehlen des Gens. Es gibt auch einen breiten Bereich von CYP2D6-Aktivitäten, zusätzlich zu schlechten Metabolisierern. Die für die Untersuchung ausgewählten Mutationen schließen C188T, G212A, delT1795, G1846T/A, G1934A, delA2637, C2938T, und G4268C ein. Die RFLPS wurden durch Gelelektrophorese nachgewiesen, wie vorher beschreiben (siehe Gough et. al., Nature 347: 773–776, 1990). Die für die RFLP verwendeten Primer entsprechen denjenigen, die durch Sachse et al, Am. J. Hum. Genet. 60: 284–295, 1997 verwendet wurden. Die für die Sequenzierung verwendeten Primer entsprechen denjenigen, die von Meyer et al., Pharmacogenetics 5: 373–384, 1995 verwendet wurden.

[0659] Das Prinzip des CMST-basierten Assays war es, ein Strang des amplifizierten CYP2D6-Exons auf einer festen Phase (z. B. magnetischen Partikeln) zu immobilisieren, die Oligonukleotidsonden zu hybridisieren, nicht hybridisiertes Material abzuwaschen, die hybridisierte Sonde zu eluieren und dann den Massenspektrometermarker durch Massenspektrometrie (nach Spalten des Markers von der Probe) nachzuweisen.

[0660] Die Amplifikationsbedingungen waren wie folgt: Primer, die das 2D6-Gen flankierten (Sachse et. al., Am. J. Hum. Genet. 60: 284–295, 1997) wurden dazu verwendet, um ein 4,681 bp genomisches DNA-Fragment zu amplifizieren, das alles der relevanten Gensequenz enthält. Die PCR-Reaktion wurde aus 1 × Expand HF Puffer, 1,5 mM MgCl<sub>2</sub>, 200 µM dNTP's, TM 0,5 µM Primern, P100 und P200, 0,5% Formamid, 100 ng gDNA, und 1,1 U Expand™ High Fidelity Enzym-Mix (Boehringer Mannheim) zusammengesetzt. Die Thermozyklusbedingungen waren wie folgt: 94°C für 3 Minuten; 10 Zyklen von 94°C für 30 Sekunden, 62°C für 30 Sekunden und 68°C für 4 Minuten, 20 Zyklen von 94°C für 30 Sekunden, 62°C für 30 Sekunden und 68°C für 4 Minuten + 20 Sekunden/Zyklus; 68°C für 10 Minuten. Die Produkte wurden auf einem 1% Agarosegel sichtbar gemacht, das mit Ethidiumbromid gefärbt wurde.

[0661] Das Assayformat wird wie folgt beschrieben: Streptavidin magnetische Partikel (Promega Magnasphere, Bindungskapazität von 80 pmol Biotin/100 µg Partikeln) wurden mit einem Niedrigsalz und Bindungspuffer gewaschen (LSWBB, 100 mM NaCl, 1 mM EDTA, 10 mM Tris, pH 7,5) und dann in einem Hochsalz Wasch-

und Bindungspuffer (HSWBB, 2 M NaCl, 1 mM EDTA, und 10 mM Tris pH 7,5) bei einer Konzentration von 2000 µg/ml resuspendiert. Die biotinylierten PCR-Produkte wurden mit den Steptavidinpartikeln für 2 Stunden bei 21°C mit konstantem Rotormischen inkubiert. Die Partikel wurden zweimal mit 200 µl HSWBB gewaschen und einmal mit 200 µl LSWBB. Die gebundenen PCR-Amplikons wurden dann durch Behandlung mit 50 µl 0,1 N NaOH für 10 Minuten bei 21°C denaturiert. Die Partikel wurden dann einmal mit 50 µl 0,1 N NaOH und dreimal mit 200 µl LSWBB gewaschen. Die Partikel-gebundenen Amplikons wurden dann mit äquimolaren Gemischen von Wild-typ (wt) und Mutanten (mt) Sonden hybridisiert, die verschiedene Massenmarker besaßen. Fünzig Picomol von jeweiliger Sonde wurden in 200 µl von 2 m GuSCN, 5 mM EDTA und 10 mM Tris pH 7,5 plaziert und 50 µl der Hybridisierungslösung wurden zu den Partikeln plaziert. Die Hybridisierung war für 1 Stunde bei 21°C mit konstantem rotierendem Mischen. Die Partikel wurden fünfmal mit LSWBB gewaschen und die Röhrchen wurden nach der zweiten Waschung gewechselt. Die hybridisierten Sonden wurden von den Partikeln durch Behandlung der Partikel mit 50 µl 0,1 N NaOH und einem anschließenden Waschschrift von 9 µl von 0,1 N NaOH eluiert. Die Lösung wurde dann mit 3 µl 1 M Essigsäure neutralisiert. 5 µl dieser Lösung wurden dann in das Massenspektrometer injiziert (HP 1100 Serie LC/MS, ausgerüstet mit Vakuum-Entgaser, binärer Pumpe, Autosampler und Diodenarraydetektor). Das Massenspektrometer wurde mit der APCI Quellenoption verwendet. HP LC/MSD Chemstation-Software wurde für die Systemkontrolle, Datenaufnahme und Datenanalyse, installiert auf einem HP vectra XA mit dem Windows NT Workstation Version 4.0 Betriebssystem) verwendet. Der Flußstrom in das MS besteht aus 50% Acetonitril in ultra-reinem Wasser bei einer Flußrate von 800 µl/Minute. Die photochemische Spaltvorrichtung bestand aus einer 254 nm Niederdruck Quecksilberlampe, einer UV-transparenten Reaktorspule und einem Lampenhalter (Aura Industries).

[0662] Repräsentative Ergebnisse sind in Tabelle 6 gezeigt.

TABELLE 6

Individuum	Exon	mAU (wt)	MAU (mt)	CMST call	RFLP call
1362 PF13	4	0	190,000	M/M	M/M
1362 PM 14	4	152,000	0	W/W	W/W
1362 MF 15	4	149,000	53,000	W/M	W/M
1377 CI 19	6	0	271,000	M/M	M/M
1377 C2 20	6	104,000	88,000	W/M	W/M
1377 C3 21	6	290,000	0	W/W	W/W
1377 CI 19	9	0	74,000	M/M	M/M
1377 C2 20	9	38,000	41,000	W/M	W/M
1377 C3 21	9	149,000	0	W/W	W/W
Kontrolle		0	0	Keines	Keines

BEISPIEL 25

## GENEXPRESSIONSVERFOLGUNG MIT CMST-MARKIERTEN ODNs

[0663] Gesamt-RNA (1–2 µg) aus der A549 menschlichen Zelllinie wurde unter der Verwendung von Superscript II reverser Transkriptase und OligodT-Primer in einem finalen Volumen von 22 µl gemäß den Anweisungen des Herstellers (Life Technologies; Gaithersburg, MD) transkribiert. Eine 75 bp-Region der Apoptose-verwandten menschliches DAD-1 Gen kodierenden Region, die eine Intron-Exon-Grenze überspannte, wurde durch Taq-Polymerasekettenreaktion (PCR) amplifiziert; eine anfängliche Denaturierung bei 95°C (5 Minuten) wurde von 25–60 Zyklen gefolgt (Annealing bei 45°C für 15 Sekunden und Denaturierung bei 95°C für 5 Sekunden). Die PCR-Reaktion (20–200 µl) enthielt 0,5 µM AD-1 reverser Primer (5'-Biotin-CCA GGA AAT TCA AAG AGT GA-3'), 0,0125 µM oder 0,5 µM DAD-1 phosphorylierter Vorwärtsprimer (5'-TTG GCT GAA TCA TTC TCA TT-3'),  $7 \times 10^2$ – $1,2 \times 10^7$  Moleküle internem Standard (5'-CC AGG AAA TTC AAA GAG TGA ACA TTC TTT TTG TGT CG-3'), 1 µl A549 cDNA oder  $8 \times 10^4$ – $10^{11}$  Moleküle von WT Mimetikum (5'-CCA GGA AAT TCA AAG AGT GAA CAT TCT TTT AGT CTC CTA CTC CTC AAT TAA GTA AAT GAG AAT GAT TCA GCC AA-3'), 0,8 U Taq-Polymerase, 0,2 mM dATP, 0,2 mM dCTP, 0,2 mM dGTP, 0,2mM dGTP, 1,5 mM MgCl<sub>2</sub>, 50

mM KCl, und 10 mM Tris-HCl pH 8,3.

[0664] Das Amplifikationsprodukt (50 µl) wurde einzelsträngig durch entweder asymmetrische Amplifikationsbedingungen (oben) oder durch Verdau mit 2,5 U Lambda Exonuklease (Boehringer Mannheim; Indianapolis, 1N) für 15 min (37°C) in 18 mM Tris-HCl pH 9,5, 1,8 mM MgCl<sub>2</sub>, 28 mM KCl gemacht. Das verdaute Amplikon oder die asymmetrische PCR-Reaktion wurden bei 70°C für 5 Min. plziert, kurz gekühlt (Raumtemperatur) und auf 2,1 M Guanidinisothiocyanate, 50 mM Tris-HCl pH 7,5, 0,1% Sarkosyl, 1 µg/ml tRNA, 5 µg/ml WT 394 Sonde (5'-394 MW CMST Marker-TTG AGG AGT AGG AGA CTA AAA-3'), 5 µg/ml IS 390 probe (5'-390 MW CMST Marker TTGACGACTACGACACAAAAA-3') eingestellt.

[0665] Die Hybridisierungsreaktion wurde für 10 Min inkubiert (RT) und auf einen 30K MWCO Spin-Filter (Millipore Corp.; Bedford, MA) transferiert, der 7 µl von 286 µg/ml Avidin-DN (Vector; Burlingame, CA) und 0,7 pg/ml tRNA enthielt. Der Spin-Filter wurde für 5–10 Min. inkubiert (RT) und zentrifugiert (4,000 × g, 10 min). Der Spin-Filter wurde zweimal mit 400 µl kaltem HPLC-Grad dH<sub>2</sub>O gewaschen und zentrifugiert (4,000 × g, 10 min). Fünfundzwanzig Mikroliters von 1 µg/ml tRNA wurden zu dem Spin-Filter hinzugefügt, um das zurückgehaltene Hybrid zu eluieren. Die 25 µl Retentat wurden in 50% Acetonitril über eine Photolyseeinheit in ein HP Massenspektrometer (APCI positiver Modus) injiziert. Einzel-Ionenmessung wurden für Marker gemacht, die Molekulargewichte von 390 und 394 aufwiesen, und die Ergebnisse als Peakbereiche dargestellt. Das Verhältnis der 390/394 Markersignale war proportional zu der Menge von Amplikon, das entweder aus dem internen Standard oder der A549 DAD-1 cDNA oder dem WT Mimetikum erzeugt wurde. Die Berechnung der unbekannten Zahl von RNA-Molekülen in der A549 Gesamt-RNA oder Zahl von WT Mimetika konnte auf einer Standardkurve gemacht werden, wie auf dem log/log-Plot der Zahl von Eingangs-internen Standardmolekülen, gegenüber dem Verhältnis der 390/394 Markersignale gezeigt. Die Unbekannte ist ein Wert von x, wenn y = 0. Alternativ ist die Unbekannte = das Verhältnis von 390/394 × der Zahl von Eingangs-interner Standardmoleküle, wenn das Verhältnis zwischen 0,3 und 3 ist.

[0666] Die Ergebnisse sind in **Fig. 23** gezeigt.

#### BEISPIEL 26

#### EINZELNUKLEOTID-VERLÄNGERUNGSSASSY

[0667] RNA-Präparation: Gesamt-RNA wurde isoliert, die aus Jurkat-Zellen (beginnend mit  $1 \times 10^8$  Zellen in exponentiellem Wachstum) unter der Verwendung eines RNA-Isolierungskits von Promega (WI) hergestellt wurde. Die RNA wurde in zwei Aliquots gelagert: (1) Vorratsaliquots in Diethylpyrocarbonat-behandeltem ddH<sub>2</sub>O wurden bei -20°C gelagert und (2) Langzeitlagerung als eine Suspension in 100% H<sub>2</sub>O.

[0668] Reverse Transkription: Poly(dT) gepimte reverse Transkription von Gesamt-RNA wurde wie beschrieben in Ausubel et al. (Ausubel et. al., in Current Protocols in Molecular Biology, 1991, Greene Publishing Associates/Wiley Interscience, NY, NY.) durchgeführt, mit der Ausnahme, daß die Reaktionen) auf die Verwendung von 1 µg von Anfangs-Gesamt-RNA eingestellt wurden. 2–50 Einheiten von reverser Transkriptase (Promega) wurden 10-fach in 10% Glycerin, 10 mM KPO<sub>4</sub>, pH 7,4, 0,2% Triton X-100, und 2 mM DTT verdünnt und auf Eis für 30 Minuten plziert, vor der Zugabe zu den Reaktionen. Gen-spezifische reverse Transkription für GAPDH und andere Kontrollgene, wie beschrieben in den Figuren und Tabellen, wurden unter Verwendung von 1 µg von Gesamt Jurkat RNA, revers transkribiert in 10 mM Tris-HCl pH 8,3, 50 mM KCl, 1 mM MgCl<sub>2</sub>, 1 mM dNTPs, 2 U/µl RNasin (Gibco-BRL), 0,1 µM Oligomer und 0,125 U/µl von M-MLV reverser Transkriptase (Gibco-BRL) in 20 µl Reaktionen durchgeführt. Die Reaktionen wurden bei 42°C für 15 Minuten inkubiert, bei 95°C für 5 Minuten Hitze-inaktiviert und auf 100 µl mit einem Mastermix von 10 mM Tris HCl pH 8,3, 1 mM NH<sub>4</sub>Cl, 1,5 M MgCl<sub>2</sub>, 100 mM KCl, 0,125 mM NTPs, 10 ng/ml der jeweiligen Oligonucleotidprimer und 0,75 Einheiten von Taq-Polymerase (Gibco-BRL) in Vorbereitung auf die PCR-Amplifikation verdünnt.

[0669] PCR: Die PCR für jedes Gen wurde mit genspezifischen Primern durchgeführt, die eine bekannte Intron/Exon-Grenze überspannten (siehe Tabellen unten). Alle PCRs wurden in 20 µl Volumina durchgeführt, die 10 mM Tris HCl pH 8,3, 1 mM NH<sub>4</sub>Cl, 1,5 M MgCl<sub>2</sub>, 100 mM KCl, 0,125 mM NTPs, 10 ng/ml der jeweiligen Oligonukleotidprimer und 0,75 Einheiten Taq-Polymerase (Gibco-BRL) enthielten. Die Zyklusparameter waren 94°C Vor-Erhitungsschritt für 5 Minuten, gefolgt von 94°C Denaturierungsschritt für 1 Minute, 55°C Annealingsschritt für 2 Minuten und einem 72°C Verlängerungsschritt für 30 Sekunden bis 1 Minute und einer finalen Verlängerung bei 72°C für 10 Minuten. Die Amplifikationen waren im allgemeinen 30–45 an der Zahl.

[0670] Reinigung von Templaten: Die PCR-Produkte wurden wie durch Zhen und Swank (Zhen und Swank, BioTechniques, 14(6): 894–898, 1993) beschrieben gelgereinigt. Die PCR-Produkte wurden auf 1% Agarosegel getrennt, die in 0,04 M Tris-Acetat, 0,001 M EDTA (1 × TAE) Puffer laufen gelassen wurden und mit Ethidiumbromid gefärbt wurden, während sie mit einer UV Lichtquelle sichtbar gemacht wurden. Eine Vertiefung wurde gerade vor der Bande von Interesse eingeschnitten und mit 50–200 µl von 10% PEG in 1 × TAE Puffer gefüllt. Die Elektrophorese wurde fortgeführt, bis die Bande vollständig in die Vertiefung eingetreten war. Die Inhalte wurden dann entfernt und extrahiert mit Phenol, Chloroform extrahiert und dann in 0,1 Volumen 7,5 M

Ammoniumacetat und 2,5 Volumen 100% EtOH gefällt. Die Proben wurden mit 75% EtOH gewaschen und kwz bei Raumtemperatur getrocknet. Die Quantifizierung der Ausbeute wurde durch Elektrophorese eines kleinen Aliquots auf 1% Agarosegel in 1 × TBE Puffer mit Ethidiumbromid-Färbung und Vergleich mit einem bekannten Standard durchgeführt.

[0671] Jede SNUPE-Reaktion wurde in einem 50 µl Volumen, das ungefähr 100 ng des amplifizierten DNA Fragments, 1 µM des SNUPE-Primers, 2 Einheiten Taq-Polymerase und 1 µl des geeigneten dNTP enthielt, durchgeführt. Alle dNTPs waren in diesem Typ Assay nicht-markiert. Der verwendete Puffer war 10 mM Tris-HCl (pH 8,3), mit 50 mM KCl, 5 mM MgCl<sub>2</sub> und 0,001% (Gew/vol) Gelatine. Die Proben wurden einem Zyklus unterzogen, der aus einer 2-Minuten Denaturierungsperiode bei 95°C, einer 2 Minuten Annealingperiode bei 60°C und einer 2-Minuten Primerverlängerungsperiode bei 72°C bestand. Die Sequenz des SNUPE-Primers für jede Familie ist unten beschrieben.

[0672] Primerverlängerungen: Einzel-Nukleotidprimerverlängerung wurden wie in Singer-Sam et al., (Singer-Sam et al., PCR Methods und Applications 1: 160–163, 1992) durchgeführt, mit der Ausnahme, daß 1 mM Mg<sup>++</sup>, 0,1 µM Primer, und 0,05 µM jedes dNTP-Typs in jedem Reaktionstyp verwendet wurden. Nach jeder oben beschriebenen Primerverlängerung wurde ein fünftel Volumen eines Ladefarbstoffs (80% Formamid, 0,1% Bromphenolblau, 0,1% Xylencyanol, 2 mM EDTA) hinzugefügt und die gesamte Probe in einem 15% denaturierenden Polyacrylamidgel elektrophoriert. Die Gele wurden in 10% Glycerin, 10% Methanol, 10% Eisessig unter konstantem Schütteln fixiert, gefolgt von Waschschritten mit 10% Glycerin. Die Gele wurden dann bei 55°C für 3–5 Stunden getrocknet.

[0673] Die in diesem Experiment beschriebenen Primer sind durch Rychlik beschrieben (Rychlik, BioTechniques 18: 84–90, 1995). Die Primer können synthetisiert werden oder als Gel-Filtrationsgrad-Primer von Midland Certified Reagent Company (Midland, Texas) erhalten werden. Die Amplifikationen sind entweder Taq DNA-Polymerase-basiert (10 mM Tris-HCl pH 8,3, 1,5 mM MgCl<sub>2</sub>, 50 mM KCl) oder Pfu-DNA-Polymerase-basiert (20 mM Tris-HCl pH 8,3, 2,0 mM MgCl<sub>2</sub>, 10 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0,1% Triton X-100, 0,1 mg/ml Rinderserumalbumin). Die Gesamtnukleosidtriphosphat (NTPs) Konzentration in den Reaktionen ist 0,8 mM, die Primerkonzentration ist 200 nM (falls nicht anders angegeben) und die Templatmenge ist 0,25 ng von Bakteriophagen Lambda-DNA pro 20 µl Reaktion. Die Zyklusparameter waren 94°C Vorheisschritt für 5 Minuten, gefolgt von 94°C Denaturierungsschritt für 1 Minute, 55°C Annealingschritt für 2 Minuten und einem 72°C Verlängerungsschritt für 30 Sekunden bis 1 Minute und einer finalen Verlängerung bei 72°C für 10 Minuten. Die Amplifikationen waren im allgemeinen 30–45 an der Zahl.

[0674] Zwei Regionen des Bakteriophagen Lambda-Genoms (GenBank Zugangsnummer # J02459) wurden als die Startstellen für die Amplifikation ausgesucht. Der 5'-Primer weist ein stabiles GC-reiches 3'-Ende auf: Der 3'-Primer ist so ausgewählt, daß sich ein 381 bp Produkt ergeben wird. Der 5'-Vorwärtsprimer ist H17: 5'-GAACGAAAACCCCCCGC. Der 3'-rückwärtige Primer ist RP 17: 5'-GATCGCCCCCAAACACATA.

[0675] Das amplifizierte Produkt wurde dann auf die Anwesenheit eines Polymorphismus an Position 31245 hin getestet. Der folgende Primer wurde in vier Einzel-Nukleotidverlängerungsassays verwendet; SNE17: 5'-GAACGAAAACCCCCCGC. Die vier Einzel-Nukleotidverlängerungsassays wurden dann wie oben beschrieben durchgeführt. All die Reaktionen wurden dann vereinigt und 5 µl des vereinigten Materials wurde auf die HPLC-Säule (SeraSep, San Jose, CA) ohne weitere Reinigung injiziert.

[0676] Die HPLC wurde unter der Verwendung von automatisierter HPLC-Instrumentierung (Rainin, Emeryville, CA. oder Hewlett Packard, Palo Alto, CA) durchgeführt. Nicht-gereinigte SNEA-Produkte, die für 3 Minuten bei 95°C vor der Injektion in eine HPLC denaturiert wurden, wurden mit einem linearen Acetonitrilgradienten (ACN, J. T. Baker, NJ) von 1,8%/Minute bei einer Flußrate von 0,9 ml/Minute eluiert. Die Start und Endpunkte wurden gemäß der Größe des SNEA-Produkts eingestellt. Die für die erfolgreiche Auflösung der SNEA-Moleküle erforderliche Temperatur war 50°C. Der Ausfluß aus der HPLC wurde dann in ein Massenspektrometer (Hewlett Packard, Palo Alto, CA) zum Nachweis von Markierungen geleitet, wobei die Ergebnisse in Tabelle 7 angegeben sind.

TABELLE 7

Markierter Primer	ddNTP-Typ	Retentionszeit	verlängert?
SNE17-487	ddATP	2,5 Minuten	nein
SNE17-496	ddGTP	2,5 Minuten	nein
SNE17-503	ddCTP	4,6 Minuten	ja
SNE17-555	ddTTP	2,5 Minuten	nein



[0677] Die Ergebnisse von Tabelle 7 zeigen, daß der Massenspektrometer-Marker (CMST) bei einer Retentionszeit von 4,6 Minuten nachgewiesen wurde, was anzeigte, daß der SNE17-Primer um eine Base (ddCTP) verlängert wurde und daher der Polymorphismus an Position 31245 in diesem Falle ein "G" war. Die SNE17-487, SNE17-496, und SNE17-555 markierten Primer wurden nicht verlängert und ihre Retentionszeiten waren auf der HPLC jeweils 2,5 Minuten.

## BEISPIEL 27

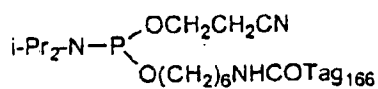
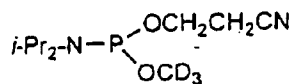
## PHOSPHORAMIDIT-CHEMIE FÜR MARKIERTE MOLEKÜLSYNTHESE

## Herstellung von Aminoethyl-verlängerten Markern

[0678] Wie in **Fig. 26** gezeigt, wurde die Herstellung von reinem Aminoethyl-verlängertem Marker (**8**) unter der Verwendung der TFP-Methodologie erhalten. Der TFP-Ester wurde durch die Reaktion des Lithiumsalz von Marker 166 (**9**) mit Überschuß-TFP-TFA und Hunigs Base in DMF präpariert. Nach einer Aufarbeitung wurde ein Dichlormethanol-TFP-Ester mit 6-Aminoethanol behandelt. Ein Präzipitat von (**8**) wurde sofort gebildet und wurde durch Filtration isoliert.

## Herstellung von Marker-Phosphoramiditen

[0679] Die Reaktion des Aminoethyl-verlängerten Markers (**5**) mit Chlorphosphoramidit ergab das gewünschte Phosphoramidit (**17**), wobei durch Massenspektrometeranalyse kein Phosphonat nachgewiesen wurde.

**(17)****(18)**

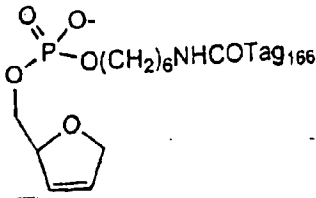
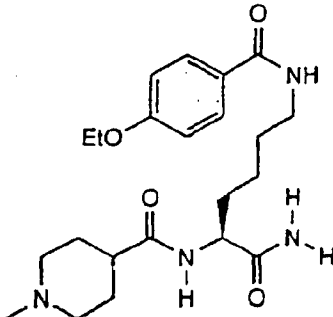
[0680] Überschüssiges phosphotylierendes Mittel wurde mit deuteriertem Methanol abgelöscht (H<sub>3</sub>-Methanol wird normalerweise verwendet, jedoch erlaubt die Verwendung von deuteriertem Methanol in diesem Fall die eindeutige Identifizierung des Ursprungs der Löschprodukte), was (**18**) hervorrief. Dieses Material überlebte die anschließende wäßrige Aufarbeitung und ist selber ein Phosphoramidit, das das Potential aufweist, in den anschließenden Reaktionen zu interferieren.

## Reaktion von Phosphoramidit mit Polymer-getragenen Thymidin

[0681] Das Thymidin wird an kontrollierte-Poren Glasperlen gekoppelt, die innerhalb einer Plastikkartusche enthalten sind. Die verwendeten Kartuschen enthalten 1000 nM ( $1 \times 10^{-6}$  Mol) von getragenen Thymidin. Das 5'-Hydroxyl der Base wird als der Dimethoxytritylether geschützt, was mit 3% Trichloressigsäure in Dichlormethan vor der Reaktion mit Phosphoramidit (**Fig. 27**) entfernt wird. Nach der Reaktion mit dem Phosphoramidit wird der Zwischenprodukt-Phosphidriester (**19**) zu dem stabileren Phosphotriester (**20**) oxidiert. Anschließendes Entfernen der Cyanoethyl-Schutzgruppe und Abspaltung von dem festen Träger wird mit Ammonium durchgeführt, um das markierte Thymidin (**21**) zu ergeben.

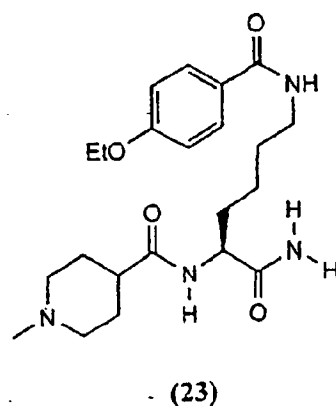
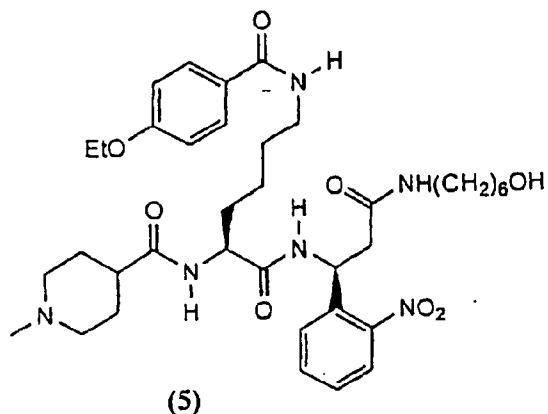
[0682] Das Rohmaterial aus der Phosphoramididherstellung wurde in dieser Sequenz verwendet und das von dem Träger abgespaltene Material durch Massenspektrometrie analysiert. Die hauptsächlichsten beobachteten Ionen im positiven Modus, zusammen mit den Zuordnungen sind in Tabelle 8 gezeigt.

TABELLE 8

Masse	Zuordnung
871	 Verlust von Thymin und Wasser
711	$\text{O}(\text{CH}_2)_6\text{NHCOTag}_{166}$ Verlust von Thymidin und Phosphat
419	 Photospaltungs-Produkt

[0683] Es scheint, daß unter den verwendeten Bedingungen die Phosphat-Esterbrücke zur Fragmentierung neigt. Die Analyse im negativen Modus zeigte einen Peak bei 338 amu, was der Verbindung (22) entspricht, dies stammt aus der Reaktion des Ablöschprodukts (18) (Fig. 28) die eine hauptsächliche Kontaminante des ungereinigten Produktes ist.

[0684] Der Marker-Phosphoramidit (17) konnte von der nicht-polaren Kontaminante (18) durch Silica-Gelchromatographie (10% Methanol/89% Dichlormethan/1% Triethylamin) abgetrennt werden und 115 mg von (17) wurden erhalten. Dieses Material wurde in dem Reaktionszyklus verwendet, um (21) zu ergeben, das als sein Spaltprodukt (5) in der Massenspektrometrie (die HPLC-Analyse des Rohprodukts zeigte, daß (5) nicht vorhanden war, daher muß der Peak bei 711 amu in der Massenspektrometrie durch die Spaltung des gewünschten Produkts (21) in dem Instrument entstanden sein) nachgewiesen wurde. Auch vorhanden war nicht-reagiertes Thymidin, was anzeigte, daß die Reaktion nicht bis zur Vervollständigung fortgeschritten war. Das Material wurde auch durch das UV-Spaltungsflußsystem hindurch passiert und das Photospaltungsprodukt (23) durch Massenspektrometrie nachgewiesen.



[0685] Daher ist es möglich, ein Massenspektrometrie-Marker-Phosphoramidit herzustellen und zu reinigen

und es mit einem Polymer-gebundenem Substrat reagieren zu lassen, das anschließen oxidiert, entschützt und von dem Träger abgespalten werden kann, um eine markierte Basis zu ergeben.

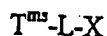
1. "Oligonucleotide Synthesis, a Practical Approach" ed M. J. Gait, IRL Press 1990.
2. R. T Pon, Tetrahedron lett, 1991, 1715.
3. Available from Glen Research Cat no 10-1953-xx.
4. "Principles of Peptide Synthesis" M. Bodansky, Springer-Verlag 1984.
5. A. J Cocuzza, Tetrahedron lett, 1989, 6287.
6. N. N. Poluskin, B. Chen, L. W. Anderson and J. S Cohen, J. Org. Chem, 1993, 58, 4606.

#### REFERENZEN

- [0686] Adams, S. P., Kavka, K. S., Wykes, E. J., Holder, S. B., and Galluppi, G. R. (1983). J. Amer. Chem. Soc. 105, 661–663.
- [0687] Andrus, A., Efcavitch, J. W., McBride, L. T., und Giusti, B. (1988). \*Tetrahedron Lett. 29, 861-864.
- [0688] Beaucage, S. L., und Caruthers, M. H. (1981). Tetrahedron Lett. 22, 1859 1862.
- [0689] Caruthers, M. H., Beaucage, S. L., Efcavitch, J. W., Fisher, E. F., Matteucci, M. D., und Stabinsky (1980). Nucleic Acids Res. Symp. Ser. 7, 215–223.
- [0690] Ferrance, I. K., Eadie, J. S., und Ivarie, R.. (1989). Nucleic Acids Res. 17, 1232–1245.
- [0691] Froehler, B. C., Ng, P. G., und Matteucci, M. D. (1986a). Nucleic Acids Res. 14, 5399–5407.
- [0692] Froehler, B. C., und Matteucci, M. D. (1986b). Tetrahedron Lett. 27, 469472.
- [0693] Froehler, B. C. (1986c). Tetrahedron Lett. 27, 5575-5578.
- [0694] Froehler, B. C., Ng, P. G., und Matteucci, M. D. (1988). Nucleic Acids Res. 16, 4831–4839.
- [0695] Gaffney, B. L. und Jones, R. A.. (1988). Tetrahedron Lett. 29, 2619–2622.
- [0696] Garegg, P. J., Regberg, T., Stawinski, J. und Stromberg, R. (1985). Chem. Scripta 25, 280–282.
- [0697] Garegg, P. J., Regberg, T., Stawinski, J. und Stromberg, R. (1986a). Chem. Scripta 26, 59–62.
- [0698] Garegg, P. J., Lidh, L., Regberg, T., Stawinski, J. und Stromberg, R. (1986b). Tetrahedron Lett. 27, 4051–4054.
- [0699] Hall, R. H., Todd, A. R., und Webb, R. F.. (1957). J. Chem. Soc., 3291–3296.
- [0700] Letsinger, R. L., und Mahadevan, V. J. (1965). J. Am. Chem. Soc. 87, 3526–3227.
- [0701] Letsinger, R. L., Finnan, J. L., Heavner, G. A. und Lunsford, W. B. (1975). J. Amer. Chem. Soc. 97, 3278.
- [0702] Letsinger, R. L., Groody, E. P., und Tanaka, T. (1982). J. Am. Chem. Soc. 104, 6805–6806.
- [0703] McBride, L. J., und Caruthers, M. H. (1983). Tetrahedron Lett. 24, 245–248.
- [0704] Merrifield, R. B. (1963). J. Am. Chem. Soc. 85, 2149–2154.
- [0705] Seliger, H. und Roesch, R. (1990). DNA und Cell Biol., 9, 691–696.
- [0706] Stein, C. A., Iversen, C. A., Subashinge, C., Cohen, J. S., Stec, W. J. und Zon, G. (1990). Analytical Biochem. 188, 11–16.
- [0707] Aus dem Voranstehenden wird es ersichtlich sein, daß, obwohl spezifische Ausführungsformen der Erfindung hier zum Zweck der Verdeutlichung beschrieben wurden, verschiedene Modifikationen gemacht werden können, ohne sich von dem Geist und Bereich der Erfindung zu entfernen.

#### Patentansprüche

1. Verbindung der Formel:

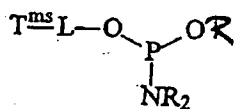


worin T<sup>ms</sup> eine organische Gruppe ist, die durch Massenspektrometrie nachweisbar ist, umfassend Kohlenstoff, mindestens eines von Wasserstoff und Fluor, und gegebenenfalls Atome, ausgewählt aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Jod;

L ist eine organische Gruppe, die es einer einmaligen T<sup>ms</sup>-enthaltenden Gruppe ermöglicht, von dem Rest der Verbindung abgespalten zu werden, wobei die T<sup>ms</sup>-enthaltende Gruppe eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterzogen wird, und die tertiäres Amin, quaternäres Amin oder organische Säure ist;

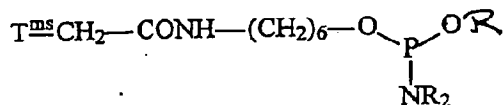
X ist eine funktionelle Gruppe, ausgewählt aus Phosphoramidit und H-Phosphonat.

2. Verbindung nach Anspruch 1, worin X eine Phosphoramiditgruppe ist, so daß T<sup>ms</sup>-L-X die Struktur



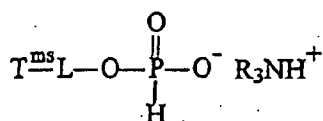
aufweist, worin R eine Alkylgruppe oder eine substituierte Alkylgruppe ist, die eine oder mehrere Substituenten ausgewählt aus Halogen und Cyano aufweist, und die zwei R-Gruppen von  $\text{NR}_2$  können miteinander verbunden sein, um eine Cycloalkylgruppe zu bilden.

3. Verbindung nach Anspruch 2, worin X eine Phosphoramiditgruppe ist, so daß  $\text{T}^{\text{ms}}\text{-L-X}$  die Struktur



aufweist und OR ist  $\text{OCR}_2\text{CH}_2\text{CN}$ , während  $\text{NR}_2$   $\text{N}(\text{iso-propyl})_2$  ist.

4. Verbindung nach Anspruch 1, worin X eine H-Phosphonatgruppe aufweist, so daß  $\text{T}^{\text{ms}}\text{-L-X}$  die Struktur



aufweist, worin R eine  $\text{C}_{1-6}$ -Alkylgruppe ist.

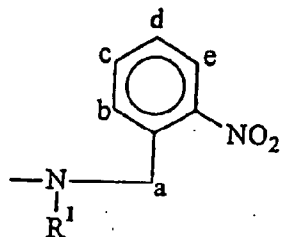
5. Verbindung nach einem der Ansprüche 1 bis 4, worin  $\text{T}^{\text{ms}}$  eine Masse von 15 bis 10.000 Dalton und eine molekulare Formel von  $\text{C}_{1-5000-100}\text{O}_{0-100}\text{S}_{0-10}\text{P}_{0-10}\text{H}_{\alpha}\text{F}_{\beta}\text{I}_{\delta}$  aufweist, worin die Summe von  $\alpha$ ,  $\beta$  und  $\delta$  ausreichend ist, die sonst nicht gesättigten Valenzen der C-, N- und O-Atome zu sättigen.

6. Verbindung nach den Ansprüchen 1 bis 4, worin  $\text{T}^{\text{ms}}$  und L miteinander über eine funktionelle Gruppe ausgewählt aus Amid, Ester, Ether, Amin, Sulfid, Thioester, Disulfid, Thioether, Harnstoff, Thioharnstoff, Carbat, Thiocarbamat, Schiffsche-Base, reduzierte Schiffsche-Base, Imin, Oxim, Hydrazon, Phosphat, Phosphonat, Phosphoramid, Phosphonamid, Sulfonat, Sulfonamid oder Kohlenstoff-Kohlenstoff-Bindung verbunden sind.

7. Verbindung nach den Ansprüchen 1 bis 4, worin L ausgewählt ist aus  $\text{L}^{\text{hv}}$ ,  $\text{L}^{\text{Säure}}$ ,  $\text{L}^{\text{Base}}$ ,  $\text{L}[\text{O}]$ ,  $\text{L}[\text{R}]$ ,  $\text{L}^{\text{enz}}$ ,  $\text{L}^{\text{elc}}$ ,  $\text{L}^{\text{A}}$  und  $\text{L}^{\text{SS}}$ , wobei jeweils aktinische Strahlung, Säure, Base, Oxidation, Reduktion, Enzym, elektrochemische, thermale und Thiol-Austausch dazu führen, daß die  $\text{T}^{\text{ms}}$ -enthaltende Gruppe von dem Rest des Moleküls abgespalten wird.

8. Verbindung nach den Ansprüchen 1 bis 4, worin  $\text{L}^{\text{hv}}$  die Formel  $\text{L}^1\text{-L}^2\text{-L}^3$  aufweist, worin  $\text{L}^2$  ein molekulares Fragment ist, daß aktinische Strahlung absorbiert, um die Abspaltung von  $\text{T}^{\text{ms}}$  von X zu unterstützen, und  $\text{L}^1$  und  $\text{L}^3$  unabhängig voneinander eine direkte Bindung oder eine organische Gruppe sind, wobei  $\text{L}^1$   $\text{L}^2$  von  $\text{T}^{\text{ms}}$  abtrennt und  $\text{L}^3$   $\text{L}^2$  von X abtrennt und weder  $\text{L}^1$  noch  $\text{L}^3$  einer Bindungsspaltung unterzogen werden, wenn  $\text{L}^2$  die actinische Strahlung absorbiert.

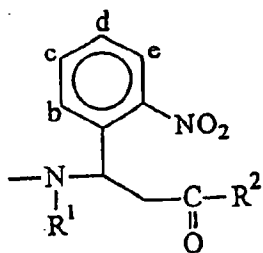
9. Verbindung nach Anspruch 8, worin  $\text{-L}^2\text{-L}^3$  die Formel:



aufweist, wobei ein Kohlenstoffatom an Positionen a, b, c, d oder e mit  $\text{-L}^3\text{-X}$  substituiert ist und gegebenenfalls eine oder mehrere der Positionen b, c, d oder e mit Alkyl, Alkoxy, Fluor, Chlor, Hydroxyl, Carboxylat oder Amid substituiert ist; und  $\text{R}^1$  Wasserstoff oder Hydrocarbyl ist.

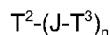
10. Verbindung nach Anspruch 8, worin  $L^3$  ausgewählt ist aus einer direkten Bindung, einem Hydrocarbylen, -O-Hydrocarbylen und Hydrocarbylen-(-O-hydrocarbylen)<sub>n</sub>-H und n eine ganze Zahl von 1 bis 10 ist.

11. Verbindung nach den Ansprüchen 1 bis 4, worin -L-X die Formel:



aufweist, worin eine oder mehrere der Positionen b, c, d oder e durch Wasserstoff, Alkyl, Alkoxy, Fluor, Chlor, Hydroxyl, Carboxylat oder Amid substituiert ist;  $R^1$  ist Wasserstoff oder Hydrocarbyl, und  $R^2$  endet in einer "X"-Gruppe.

12. Verbindung nach den Ansprüchen 1 bis 4, worin  $T^{ms}$  die Formel:



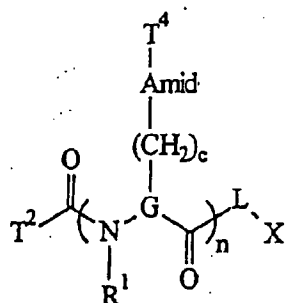
aufweist,  $T^2$  eine organische Gruppe ist, die aus Kohlenstoff und einem oder mehreren von Wasserstoff, Fluor, Jod, Sauerstoff, Stickstoff, Schwefel und Phosphor gebildet ist, die eine Masse von 15 bis 500 Dalton aufweist;  $T^3$  eine organische Gruppe ist, die aus Kohlenstoff und einem oder mehreren von Wasserstoff, Fluor, Jod, Sauerstoff, Stickstoff, Schwefel und Phosphor gebildet ist, und die eine Masse von 50 bis 1.000 Dalton aufweist; J ist eine direkte Verbindung oder eine funktionelle Gruppe, ausgewählt aus Amid, Ester, Amin, Sulfid, Ether, Thioester, Disulfid, Thioether, Harnstoff, Thioharnstoff, Carbamat, Thiocarbamat, Schiffsche-Base, reduzierte Schiffsche-Base, Imin, Oxim, Hydrazon, Phosphat, Phosphonat, Phosphoramid, Phosphonamid, Sulfonat, Sulfonamid oder Kohlenstoff-Kohlenstoff-Bindung; und n ist eine ganze Zahl von 1 bis 50, und wenn n größer als 1 ist, ist jedes von  $T^3$  und J unabhängig voneinander ausgewählt.

13. Verbindung nach Anspruch 12, worin  $T^2$  ausgewählt ist aus Hydrocarbyl, Hydrocarbyl-O-hydrocarbylen, Hydrocarbyl-S-hydrocarbylen, Hydrocarbyl-NH-hydrocarbylen, Hydrocarbyl-amid-hydrocarbylen, N-(Hydrocarbyl)hydrocarbylen, N,N-Di(hydrocarbyl)hydrocarbylen, Hydrocarbylacyl-hydrocarbylen, Heterocyclylhydrocarbyl, worin das/die Heteroatom(e) ausgewählt ist/sind aus Sauerstoff, Stickstoff, Schwefel und Phosphor, substituiertes Heterocyclylhydrocarbyl, worin das/die Heteroatom(e) ausgewählt ist/sind aus Sauerstoff, Stickstoff, Schwefel und Phosphor und die Substituenten ausgewählt sind aus Hydrocarbyl, Hydrocarbyl-O-hydrocarbylen, Hydrocarbyl-NH-hydrocarbylen, Hydrocarbyl-S-hydrocarbylen, N-(Hydrocarbyl)hydrocarbylen, N,N-Di(hydrocarbyl)hydrocarbylen und Hydrocarbylacyl-hydrocarbylen, sowie Derivative von jedem der Voranstehenden, wobei einer oder mehrere Waerstoffe durch eine gleiche Zahl von Fluor ersetzt ist.

14. Verbindung nach Anspruch 12, worin  $T^3$  die Formel -G( $R^2$ )- aufweist, G ist  $C_{1-6}$  Alkyl mit einem einzelnen  $R^2$  Substituenten, und  $R^2$  ist ausgewählt aus Alkyl, Alkenyl, Alkynyl, Cycloalkyl, Aryl-fusioniertes-Cycloalkyl, Cycloalkenyl, Aryl, Aralkyl, Arylsubstituiertes Alkenyl oder Alkynyl, Cycloalkyl-substituiertes Alkyl, Cycloalkenylsubstituiertes Cycloalkyl, Biaryl, Alkoxy, Alkenoxy, Alkinoxy, Aralkoxy, Arylsubstituiertes Alkenoxy oder Alkinoxy, Alkylamino, Alkenylamino oder Alkynylamino, Aryl-substituiertes Alkylamino, Aryl-substituiertes Alkenylamino oder Alkynylamino, Aryloxy, Arylamino, N-Alkylharnstoff-substituiertes Alkyl, N-arylharnstoff-substituiertes Alkyl, Alkylcarbonylamino-substituiertes Alkyl, Aminocarbonyl-substituiertes Alkyl, Heterocyclyl, Heterocyclyl-substituiertes Alkyl, Heterocyclyl-substituiertes Amino, Carboxyalkyl substituiertes Aralkyl, Oxocarbocyclylfusioniertes Aryl und Heterocyclylalkyl; Cycloalkenyl, Aryl-substituiertes Alkyl und Aralkyl, Hydroxy-substituiertes Alkyl, Alkoxy-substituiertes Alkyl, Aralkoxy-substituiertes Alkyl, Alkoxy-substituiertes Alkyl, Aralkoxy-substituiertes Alkyl, Amino-substituiertes Alkyl, (Aryl-substituiertes Alkyloxycarbonylamino)-substituiertes Alkyl, Thiol-substituiertes Alkyl, Alkylsulfonyl-substituiertes Alkyl, (Hydroxy-substituiertes Alkylthio)-substituiertes Alkyl, Thioalkoxy-substituiertes Alkyl, Hydrocarbylacylamino-substituiertes Alkyl, Heterocyclylacylamino-substituiertes Alkyl, Hydrocarbyl-substituiertes-Heterocyclylacylamino-substituiertes Alkyl, Alkylsulfonylamino-substituiertes Alkyl, Arylsulfonylamino-substituiertes Alkyl, Morpholino-alkyl, Thiomorpholino-alkyl, Morpholinocarbonyl-substituiertes Alkyl, [N-(Alkyl, Alkenyl or Alkynyl)- oder N,N-[Dialkyl, Dialkenyl, Dialkynyl oder (Alkyl, Alkenyl)-amino]carbonyl-substituiertes Alkyl, Heterocyclyla-

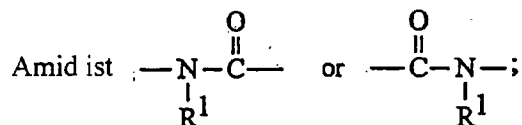
minocarbonyl, Heterocylylalkylenaminocarbonyl, Heterocyclylaminocarbonyl-substituiertes Alkyl, Heterocylylalkylenaminocarbonyl-substituiertes Alkyl, N,N-[Dialkyl]alkylenaminocarbonyl, N,N-[Dialkyl]alkylenaminocarbonyl-substituiertes Alkyl, Alkyl-substituiertes Heterocyclylcarbonyl, Alkyl-substituiertes Heterocyclylcarbonyl-alkyl, Carboxyl-substituiertes Alkyl, Dialkylamino-substituiertes Acylaminoalkyl und Aminosäure-Seitenketten, ausgewählt aus Arginin, Asparagin, Glutamin, S-Methylcystein, Methionin und entsprechenden Sulfoxid- und Sulfonderivaten davon, Glycin, Leucin, Isoleucin, Allo-Isoleucin, Tert-Leucin, Norleucin, Phenylalanin, Tyrosin, Tryptophan, Prolin, Alanin, Ornithin, Histidin, Glutamin, Valin, Threonin, Serin, Asparaginsäure, beta-Cyanoalanin, und Allo-Threonin; Alinyl- und Heterocyclylcarbonyl, Aminocarbonyl, Amido, Mono- oder Dialkylaminocarbonyl, Mono- oder Diarylaminocarbonyl, Alkylarylaminocarbonyl, Diarylaminocarbonyl, Mono- oder Diacylaminocarbonyl, aromatischem oder aliphatischem Acyl, Alkyl, gegebenenfalls substituiert durch Substituenten, ausgewählt aus Amino, Carboxy, Hydroxy, Mercapto, Mono- oder Dialkylamino, Mono- oder Diarylamino, Alkylarylamino, Diarylamino, Mono- oder Diacylamino, Alkoxy, Alkenoxy, Aryloxy, Thioalkoxy, Thioalkenoxy, Thioalkinoxy, Thioaryloxy und Heterocyclyl.

15. Verbindung nach Anspruch 12, die die folgende Formel:



aufweist, worin G  $(CH_2)_{1-6}$  ist, worin ein Wasserstoff an einer und nur einer der  $CH_2$ -Gruppen durch  $-(CH_2)_c$ -Amid- $T^4$  ersetzt ist;

$T^2$  und  $T^4$  organische Gruppen der Formel  $C_{1-25}N_{0-9}O_{0-9}H_{\alpha}F_{\beta}$  sind, worin die Summe von  $\alpha$  und  $\beta$  ausreichend ist, die sonst ungesättigten Valenzen von C-, N- und O-Atomen zu sättigen;



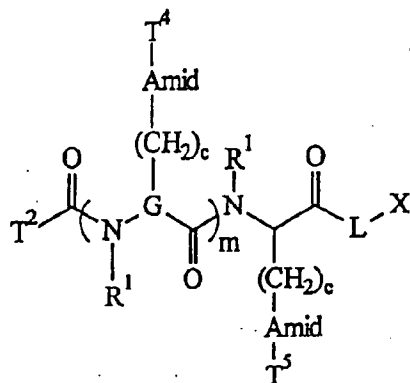
$R^1$  ist Wasserstoff oder  $C_{1-10}$  Alkyl;

c ist eine ganze Zahl im Bereich von 0-4;

X ist gemäß Anspruch 1 definiert; und

n ist eine ganze Zahl im Bereich von 1 bis 50, so daß, wenn n größer als 1 ist, G, c, Amid,  $R^1$  und  $T^4$  unabhängig voneinander ausgewählt sind.

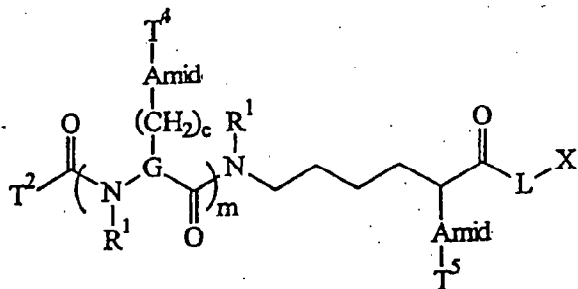
16. Verbindung nach Anspruch 15, die die Formel



aufweist, worin  $T^5$  eine organische Gruppe der Formel  $C_{1-25}N_{0-9}O_{0-9}H_{\alpha}F_{\beta}$  ist, worin die Summe von  $\alpha$  und  $\beta$  ausreichend ist, die sonst ungesättigten Valenzen der C-, N- und O-Atome zu sättigen und  $T^5$  ein tertiäres oder

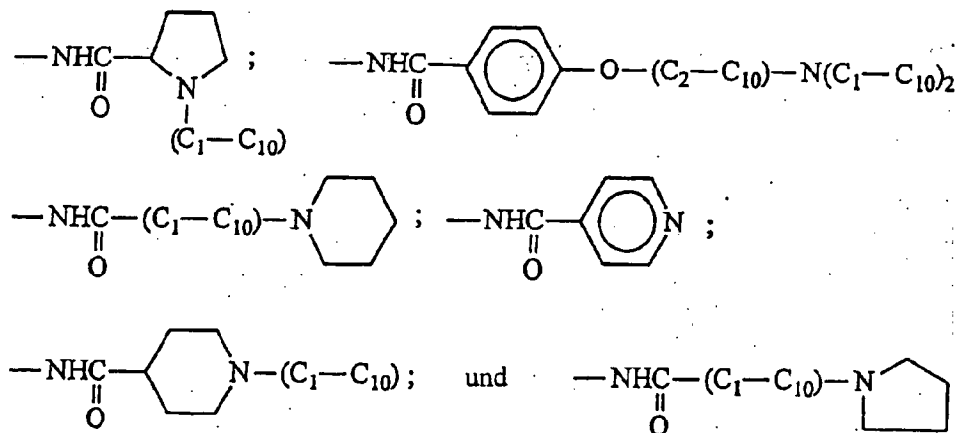
quaternäres Amin oder eine organische Säure einschließt; und m eine ganze Zahl im Bereich von 0–49 ist.

17. Verbindung nach Anspruch 15, die die Formel

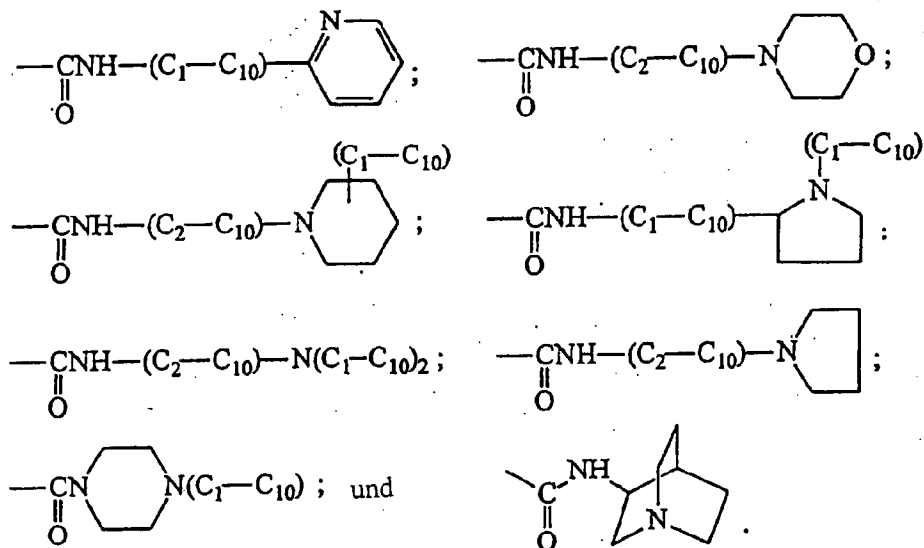


aufweist, worin  $T^5$  eine organische Gruppe der Formel  $C_{1-25}N_{0-9}4_{0-9}H_{\alpha}F_{\beta}$  ist, worin die Summe von  $\alpha$  und  $\beta$  ausreichend ist, um die sonst ungesättigten Valenzen der C-, N- und O-Atome zu sättigen; und  $T^5$  ein tertiäres oder quaternäres Amin oder eine organische Säure einschließt; und m eine ganze Zahl im Bereich von 0–49 ist.

18. Verbindung nach einem der Ansprüche 16 und 17, worin -Amid- $T^5$  ausgewählt ist aus:



19. Verbindung nach einem der Ansprüche 16 und 17, worin -Amid- $T^5$  ausgewählt ist aus:



20. Verbindung nach einem der Ansprüche 14–16, worin  $T^2$  die Struktur aufweist, die sich ergibt, wenn eine der folgenden organischen Säuren mit einer Amin-Gruppe kondensiert wird, um  $T^2-C(=O)-NCR^1$ - zu bilden: Ameisensäure, Essigsäure, Propiolsäure, Propionsäure, Fluoressigsäure, 2-Butionsäure, Cyclopropancar-

boxylsäure, Buttersäure, Methoxyessigsäure, Difluoressigsäure, 4-Pentinsäure, Cyclobutanocarboxylsäure, 3,3-Dimethylacrylsäure, Valerinsäure, N,N-Dimethylglycin, N-Formyl-Gly-OH, Ethoxyessigsäure, (Methylthio)essigsäure, Pyrrol-2-carboxylsäure, 3-Furoesäure, Isoxazol-5-carboxylsäure, trans-3-Hexenonsäure, Trifluoressigsäure, Hexanonsäure, Ac-Gly-OH, 2-Hydroxy-2-methylbuttersäure, Benzoesäure, Nicotinsäure, 2-Pyrazincarboxylsäure, 1-Methyl-2-pyrrolcarboxylsäure, 2-Cyclopenten-1-essigsäure, Cyclopentylessigsäure, (S)-(-)-2-Pyrrolidon-5-carboxylsäure, N-Methyl-L-prolin, Heptanonsäure, Ac-b-Ala-OH, 2-Ethyl-2-hydroxybuttersäure, 2-(2-Methoxyethoxy)essigsäure, p-Toluonsäure, 6-Methylnicotinsäure, 5-Methyl-2-pyrazincarboxylsäure, 2,5-Dimethylpynol-3-carboxylsäure, 4-Fluorbenzoesäure, 3,5-Dimethylisoxazol-4-carboxylsäure, 3-Cyclopentylpropionsäure, Octanonsäure, N,N-Dimethylbernsteinsäure, Phenylpropioninsäure, Zimtsäure, 4-Ethylbenzoesäure, p-Anisinsäure, 1,2,5-Trimethylpynol-3-carboxylsäure, 3-Fluor-4-methylbenzoesäure, Ac-DL-Propargylglycin, 3-(Trifluormethyl)buttersäure, 1-Piperidinpropionsäure, N-Acetylprolin, 3,5-Difluorbenzoesäure, Ac-L-Val-OH, Indol-2-carboxylsäure, 2-Benzofurancarboxylsäure, Benzotriazol-5-carboxylsäure, 4-n-Propylbenzoesäure, 3-Dimethylaminobenzoessäure, 4-Ethoxybenzoesäure, 4-(Methylthio)benzoesäure, N-(2-Furoyl)glycin, 2-(Methylthio)nicotinsäure, 3-Fluor-4-methoxybenzoesäure, Tfa-Gly-OH, 2-Napthoinsäure, Chinaldinsäure, Ac-L-Ile-OH, 3-Methylinden-2-carboxylsäure, 2-Chinoxalincarboxylsäure, 1-Methylindenol-2-carboxylsäure, 2,3,6-Trifluorbenzoesäure, N-Formyl-L-Met-OH, 2-[2-(2-Methoxyethoxy)ethoxy]essigsäure, 4-n-Butylbenzoesäure, N-Benzoylglycin, 5-Fluorindol-2-carboxylsäure, 4-n-Propoxybenzoesäure, 4-Acetyl-3,5-dimethyl-2-pyrrolcarboxylsäure, 3,5-Dimethoxybenzoesäure, 2,6-Dimethoxynicotinsäure, Cyclohexanpentanonsäure, 2-Naphthylessigsäure, 4-(1H-Pyrrol-1-yl)benzoesäure, Indol-3-propionsäure, m-Trifluormethylbenzoesäure, 5-Methoxyindol-2-carboxylsäure, 4-Pentylbenzoesäure, Bz-b-Ala-OH, 4-Diethylaminobenzoessäure, 4-n-Butoxybenzoesäure, 3-Methyl-5-CF<sub>3</sub>-isoxazol-4-carboxylsäure, (3,4-Dimethoxyphenyl)essigsäure, 4-Biphenylcarboxylsäure, Pivaloyl-Pro-OH, Octanoyl-Gly-OH, (2-Naphthoxy)essigsäure, Indol-3-buttersäure, 4-(Trifluormethyl)phenylessigsäure, 5-Methoxyindol-3-essigsäure, 4-Trifluormethoxybenzoesäure, Ac-L-Phe-OH, 4-Pentyloxybenzoesäure, Z-Gly-OH, 4-Carboxy-N-(fur-2-ylmethyl)pynolidin-2-on, 3,4-Diethoxybenzoesäure, 2,4-Dimethyl-5-CO<sub>2</sub>Et-pyrrol-3-carboxylsäure, N-(2-Fluorphenyl)bernsteinsäure, 3,4,5-Trimethoxybenzoesäure, N-Phenylanthranilinsäure, 3-Phenoxybenzoesäure, Nonanoyl-Gly-OH, 2-Phenoxy-pyridin-3-carboxylsäure, 2,5-Dimethyl-1-phenylpynol-3-carboxylsäure, trans-4-(Trifluormethyl)zinn-säure, (5-Methyl-2-phenyloxazol-4-yl)essigsäure, 4-(2-Cyclohexenyloxy)benzoesäure, 5-Methoxy-2-methylindenol-3-essigsäure, trans-4-Cotinincarboxylsäure, Bz-5-Aminovalerinsäure, 4-Hexyloxybenzoesäure, N-(3-Methoxyphenyl)bernsteinsäure, Z-Sar-OH, 4-(3,4-Dimethoxyphenyl)buttersäure, Ac-o-Fluor-DL-Phe-OH, N-(4-Fluorphenyl)glutarminsäure, 4'-Ethyl-4-biphenylcarboxylsäure, 1,2,3,4-Tetrahydroacridincarboxylsäure, 3-Phenoxyphenylessigsäure, N-(2,4-Difluorphenyl)bernsteinsäure, N-Decanoyl-Gly-OH, (+)-6-Methoxy-a-methyl-2-naphthalinessigsäure, 3-(Trifluormethoxy)zinnaminsäure, N-Formyl-DL-Trp-OH, (R)-(+)-a-Methoxy-a-(trifluormethyl)phenylessigsäure, Bz-DL-Leu-OH, 4-(Trifluormethoxy)phenoxyessigsäure, 4-Heptyloxybenzoesäure, 2,3,4-Trimethoxyzinnaminsäure, 2,6-Dimethoxybenzoyl-Gly-OH, 3-(3,4,5-Trimethoxyphenyl)propionsäure, 2,3,4,5,6-Pentafluorphenoxyessigsäure, N-(2,4-Difluorphenyl)glutarminsäure, N-Undecanoyl-Gly-OH, 2-(4-Fluorbenzoyl)benzoesäure, 5-Trifluormethoxyindol-2-carboxylsäure, N-(2,4-Difluorphenyl)diglycolaminsäure, Ac-L-Trp-OH, Tfa-L-Phenylglycin-OH, 3-Jodbenzoesäure, 3-(4-n-Pentylbenzoyl)propionsäure, 2-Phenyl-4-chinolinicarboxylsäure, 4-Octyloxybenzoesäure, Bz-L-Met-OH, 3,4,5-Triethoxybenzoesäure, N-Lauroyl-Gly-OH, 3,5-Bis(trifluormethyl)benzoesäure, Ac-5-Methyl-DL-Tip-OH, 2-Jodphenylessigsäure, 3-Jod-4-methylbenzoesäure, 3-(4-n-Hexylbenzoyl)propionsäure, N-Hexanoyl-L-Phe-OH, 4-Nonyloxybenzoesäure, 4'-(Trifluormethyl)-2-biphenylcarboxylsäure, Bz-L-Phe-OH, N-Tridecanoyl-Gly-OH, 3,5-Bis(trifluormethyl)phenylessigsäure, 3-(4-n-Heptylbenzoyl)propionsäure, N-Hepytanoyl-L-Phe-OH, 4-Decyloxybenzoesäure, N-( $\alpha,\alpha,\alpha$ -Trifluor-m-tolyl)anthranilinsäure, Nifluminsäure, 4-(2-Hydroxyhexafluorisopropyl)benzoesäure, N-Myristoyl-Gly-OH, 3-(4-n-Octylbenzoyl)propionsäure, N-Octanoyl-L-Phe-OH, 4-Undecyloxybenzoesäure, 3-(3,4,5-Trimethoxyphenyl)propionyl-Gly-OH, 8-Jodnaphthoinsäure, N-Pentadecanoyl-Gly-OH, 4-Dodecyloxybenzoesäure, N-Palmitoyl-Gly-OH und N-Stearoyl-Gly-OH.

21. Verfahren zur Bestimmung der Anwesenheit eines Single Nukleotide Polymorphismus in einem Nukleinsäureziel, umfassend:

- a) Amplifizieren einer Sequenz eines Nukleinsäureziels, das einen Single Nukleotide Polymorphismus enthält;
- b) Erzeugen einer Einzelstrangform des Ziels;
- c) Kombinieren einer markierten Nukleinsäuresonde mit den amplifizierten Zielnukleinsäuremolekülen unter Bedingungen und für eine Zeitdauer, die ausreichend ist, um eine Hybridisierung der markierten Nukleinsäuresonde an komplementäre amplifizierte ausgewählte Zielnukleinsäuremoleküle zu ermöglichen, wobei die Markierung mit einem bestimmten Single Nukleotide Polymorphismus korrelierbar ist und durch Spektrometrie oder Potentiometrie nachweisbar ist;
- d) Abtrennen von nicht hybridisierter markierter Probe von hybridisierter markierter Probe durch eine Größenbestimmungs-Methodologie;
- e) Abspalten des Markers von der Sonde; und



f) Nachweisen des Markers durch Spektrometrie oder Potentiometrie und Bestimmen der Anwesenheit des Single Nukleotide Polymorphismus, wobei die markierte Nukleinsäuresonde die Struktur

$T^{ms}-L-X'$

aufweist, worin  $T^{ms}$  eine organische Gruppe ist, die durch Massenspektrometrie nachweisbar ist, umfassend Kohlenstoff, mindestens eines von Wasserstoff und Fluor, und gegebenenfalls Atome ausgewählt aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Jod;

L ist eine organische Gruppe, die es einer einmaligen  $T^{ms}$ -enthaltenden Gruppe ermöglicht, von dem Rest der Verbindung abgespalten zu werden, wobei die  $T^{ms}$ -enthaltende Gruppe eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterzogen wird, und die tertiäres Amin, quaternäres Amin oder organische Säure ist; und

$X'$  ist Nukleinsäure.

22. Verfahren zur Bestimmung der Anwesenheit eines Single Nukleotide Polymorphismus in einem Nukleinsäureziel, umfassend:

- Amplifizieren einer Sequenz eines Nukleinsäureziels, das einen Single Nukleotide Polymorphismus enthält;
- Kombinieren eines markierten Nukleinsäureprimers mit den amplifizierten Zielnukleinsäuremolekülen unter Bedingungen und für eine Zeitdauer, die ausreichend ist, um ein Annealing des markierten Nukleinsäureprimers an komplementäre amplifizierte ausgewählte Zielnukleinsäuremoleküle zu ermöglichen, wobei der Oligonukleotidprimer eine zu der Wildtyp-Sequenz oder dem Single Nukleotide Polymorphismus komplementäre äußerst 3'-liegende Base aufweist, wobei der Marker mit einem bestimmten Single Nukleotide Polymorphismus korrelierbar ist und durch Spektrometrie oder Potentiometrie nachweisbar ist;
- Verlängern des Primers, wobei ein zu dem Ziel komplementärer Strang synthetisiert wird, wenn die am weitesten 3'-liegende Base des Primers mit dem Ziel komplementär ist;
- Abtrennen von nicht verlängertem markierten Primer von verlängertem markierten Primer durch eine Größenbestimmungs-Methodologie;
- Abspalten des Markers von den Primern oder verlängerten Primern; und
- Nachweisen des Markers durch Spektrometrie oder Potentiometrie, und Bestimmen daraus der Anwesenheit des Single Nukleotide Polymorphismus, wobei der markierte Nukleinsäureprimer die Struktur

$T^{ms}-L-X$

aufweist, worin  $T^{ms}$  eine organische Gruppe ist, die durch Massenspektrometrie nachweisbar ist, umfassend Kohlenstoff, mindestens eines von Wasserstoff und Fluorid, und gegebenenfalls Atome ausgewählt aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Jod;

L ist eine organische Gruppe, die es einer einmaligen  $T^{ms}$ -enthaltenden Gruppe ermöglicht, von dem Rest der Verbindung abgespalten zu werden, wobei die  $T^{ms}$ -enthaltende Gruppe eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterzogen wird, und die tertiäres Amin, quaternäres Amin oder organische Säure ist; und

$X'$  ist Nukleinsäure.

23. Verfahren zur Bestimmung der Menge eines spezifischen mRNA-Moleküls in einer Nukleinsäurepopulation, umfassend:

- Überführen einer RNA-Population in eine cDNA-Population;
- Hinzufügen einer einzelsträngigen Nukleinsäure (interner Standard), die eine Vielzahl von Single Nukleotide Polymorphismen enthält, die ansonsten identisch zu dem cDNA-Ziel ist;
- Amplifizieren einer spezifischen Sequenz des cDNA-Ziels;
- Coamplifizieren des internen Standards, wobei der interne Standard dieselbe Länge aufweist, wie das cDNA-Amplicon;
- Erzeugen einer einzelsträngigen Form des Ziels;
- Kombinieren eines Sets von markierten Nukleinsäuresonden mit der amplifizierten Ziel-cDNA und dem amplifizierten internen Standard unter Bedingungen für eine Zeitdauer, die ausreichend ist, um eine Hybridisierung der markierten Nukleinsäuresonde an komplementäre ausgewählte Ziel-cDNA und interne Standardsequenzen zu erlauben, wobei der Marker mit einer bestimmten cDNA-Sequenz korrelierbar ist und ein zweiter Marker ist mit dem internen Standard korrelierbar, und ist durch Spektrometrie oder Potentiometrie nachweisbar;
- Abtrennen von nicht hybridisierter markierter Sonde von hybridisierter markierter Sonde durch eine Größenbestimmungs-Methodologie;

- h) Abspalten des Markers von den Sonden;
- i) Nachweisen der Marker durch Spektrometrie oder Potentiometrie; und
- j) Nehmen des Verhältnisses von Marker, der zu cDNA koneliert ist, gegenüber Marker, der mit dem internen Standard koneliert ist und daraus bestimmen der Menge der cDNA, wodurch die Menge der spezifischen mRNA in einer Nukleinsäurepopulation bestimmt wird, wobei die markierte Nukleinsäuresonde die Struktur

$T^{ms}-L-X'$

aufweist, worin  $T^{ms}$  eine organische Gruppe ist, die durch Massenspektrometrie nachweisbar ist, umfassend Kohlenstoff, mindestens eines von Wasserstoff und Fluor, und gegebenenfalls Atome ausgewählt aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Jod;

L ist eine organische Gruppe, die es einer einmaligen  $T^{ms}$ -enthaltenden Gruppe ermöglicht, von dem Rest der Verbindung abgespalten zu werden, wobei die  $T^{ms}$ -enthaltende Gruppe eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterzogen wird, und die tertiäres Amin, quaternäres Amin oder organische Säure ist; und

$X'$  ist Nukleinsäure.

24. Verfahren zur Bestimmung der Menge eines Single Nukleotide Polymorphismus in einem Nukleinsäureziel, umfassend:

- a) Amplifizieren einer Sequenz eines Nukleinsäureziels, das einen Single Nukleotide Polymorphismus aufweist;
- b) Erzeugen einer einzelsträngigen Form des Ziels;
- c) Kombinieren einer markierten Nukleinsäuresonde mit den amplifizierten Zielnukleinsäuremolekülen unter Bedingungen und für eine Zeitdauer, die ausreichend ist, um eine Hybridisierung der markierten Nukleinsäuresonde an komplementäre amplifizierte ausgewählte Zielnukleinsäuremoleküle zu ermöglichen, wobei der Marker mit einem bestimmten Single Nukleotide Polymorphismus korrelierbar ist und durch Spektrometrie oder Potentiometrie nachweisbar ist;
- d) Abtrennen von nicht hybridisierter markierter Sonde von markierter hybridisierter Sonde durch eine Größenbestimmungs-Methodologie;
- e) Abspalten des Markers von den Sonden;
- f) Nachweisen der Marker durch Spektrometrie oder Potentiometrie; und
- g) Nehmen des Verhältnisses von mit dem Wildtyp-Polymorphismus korrelierbaren Markers zu dem mit dem Mutantenpolymorphismus korrelierten Markers und daraus bestimmen der Menge des Polymorphismus, wobei die markierte Nukleinsäure die Struktur

$T^{ms}-L-X'$

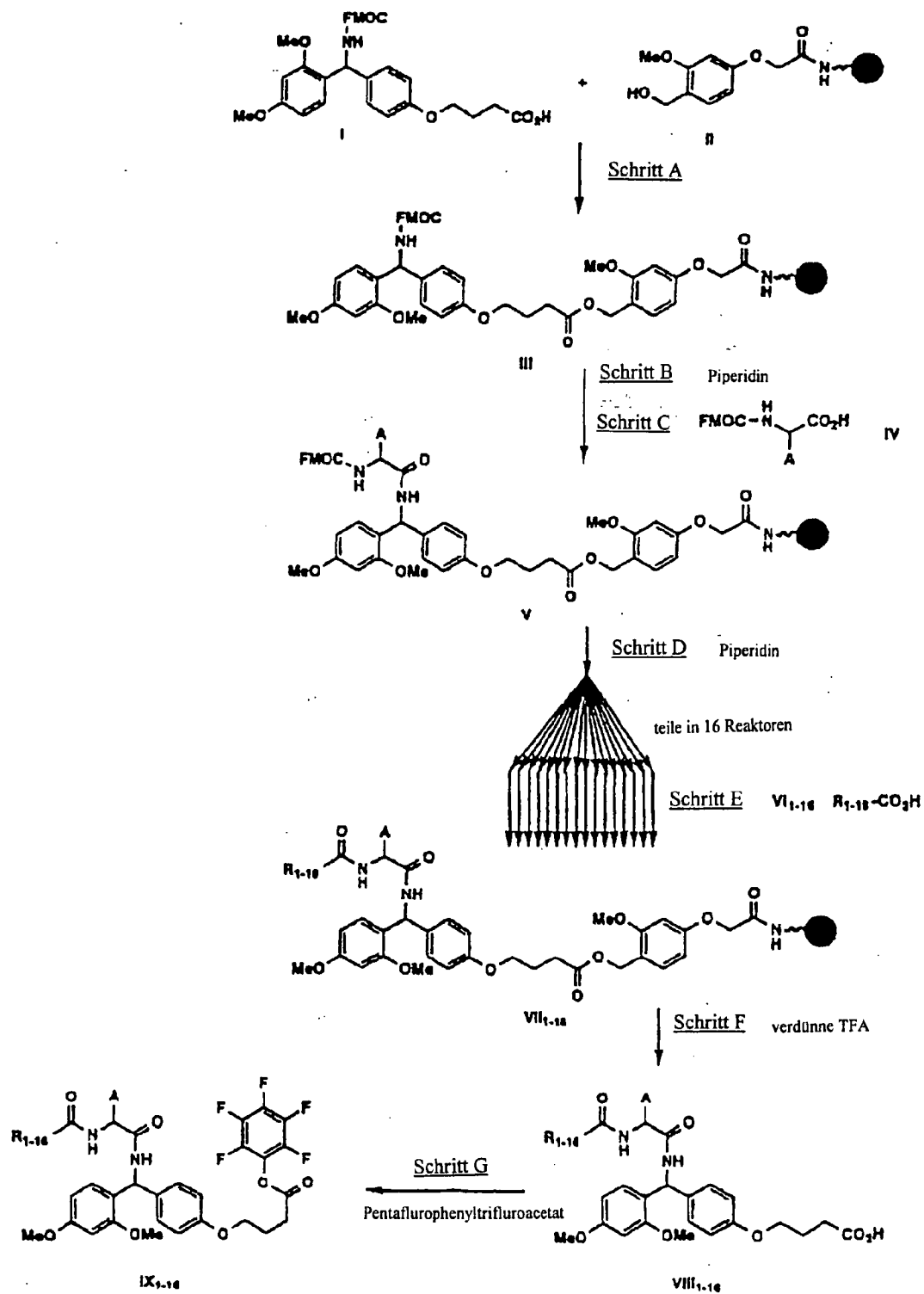
aufweist, worin  $T^{ms}$  eine organische Gruppe ist, die durch Massenspektrometrie nachweisbar ist, umfassend Kohlenstoff, mindestens eines von Wasserstoff und Fluor, und gegebenenfalls Atome ausgewählt aus Sauerstoff, Stickstoff, Schwefel, Phosphor und Jod;

L ist eine organische Gruppe, die es einer einmaligen  $T^{ms}$ -enthaltenden Gruppe ermöglicht, von dem Rest der Verbindung abgespalten zu werden, wobei die  $T^{ms}$ -enthaltende Gruppe eine funktionelle Gruppe umfaßt, die einen einzelnen ionisierten Ladungszustand unterstützt, wenn die Verbindung Massenspektrometrie unterzogen wird, und die tertiäres Amin, quaternäres Amin oder organische Säure ist; und

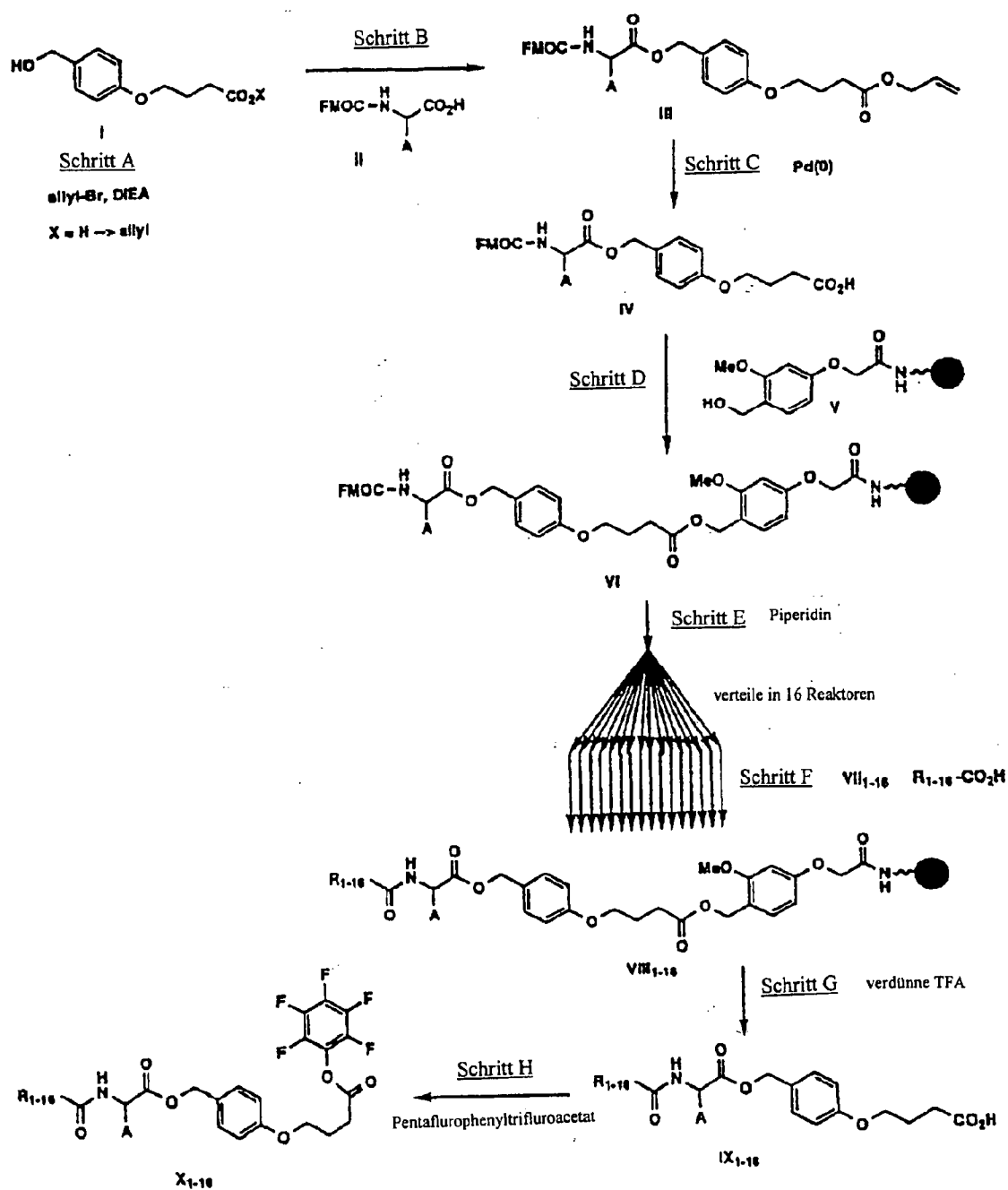
$X'$  ist Nukleinsäure.

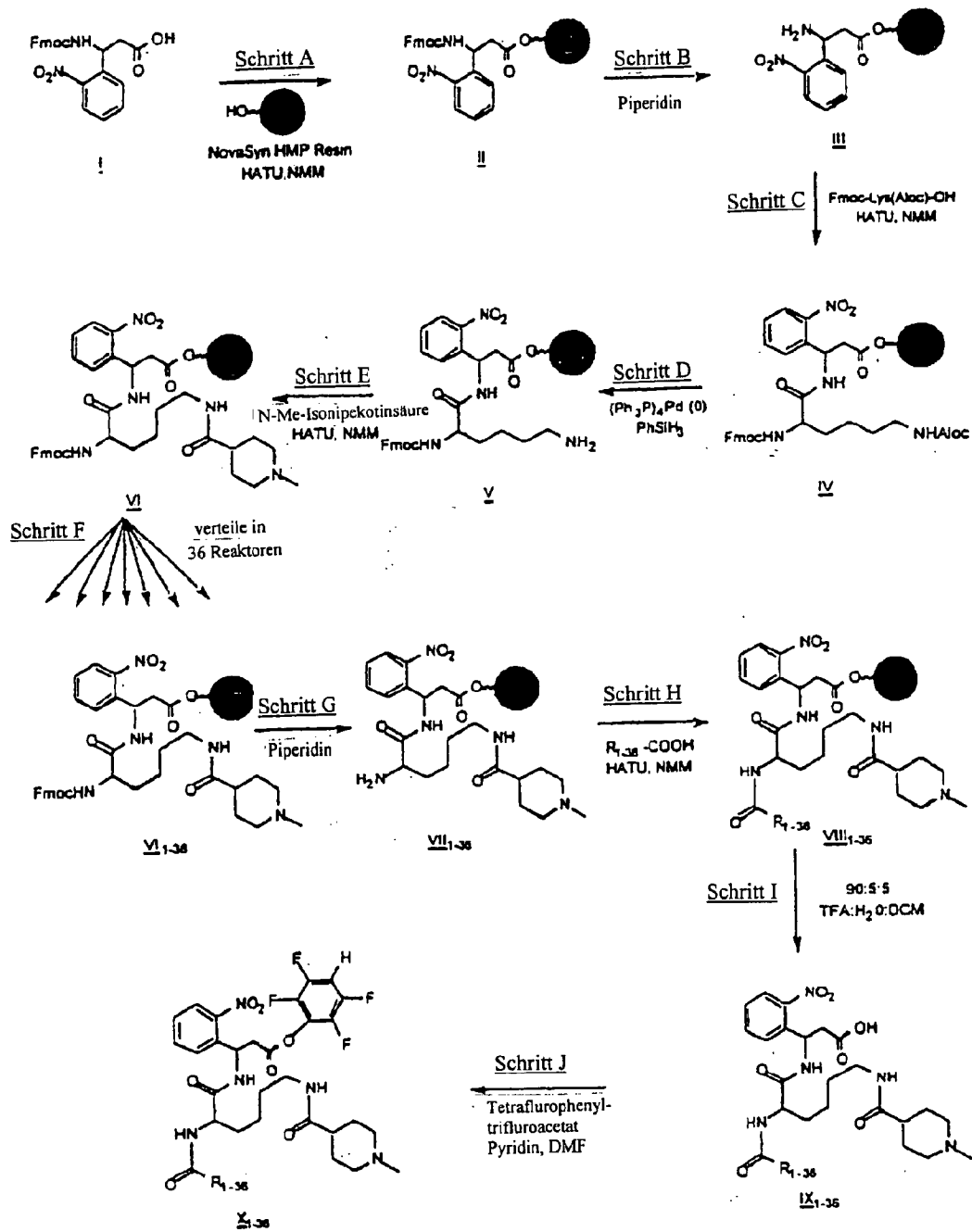
Es folgen 36 Blatt Zeichnungen

FIGUR 1

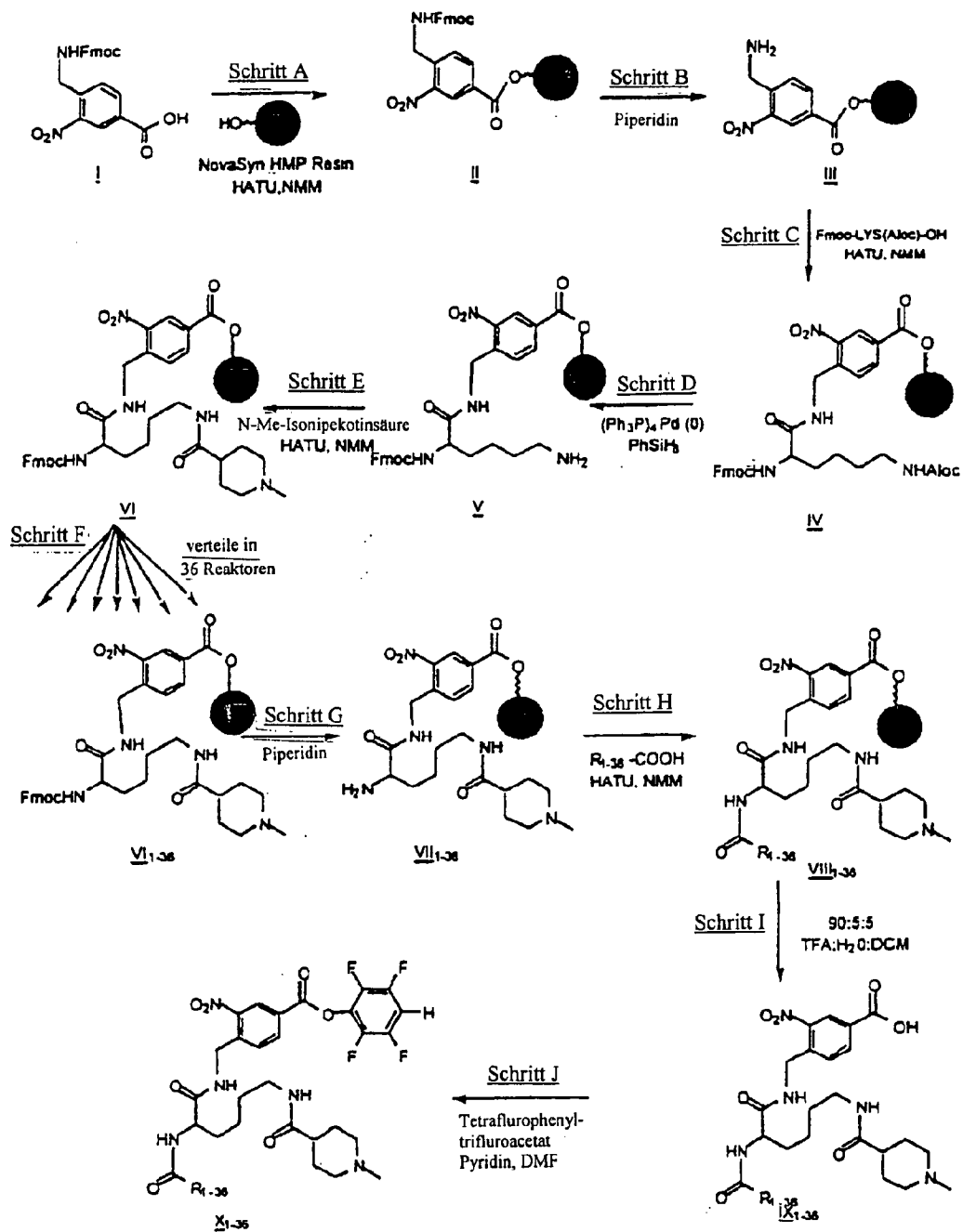


FIGUR 2

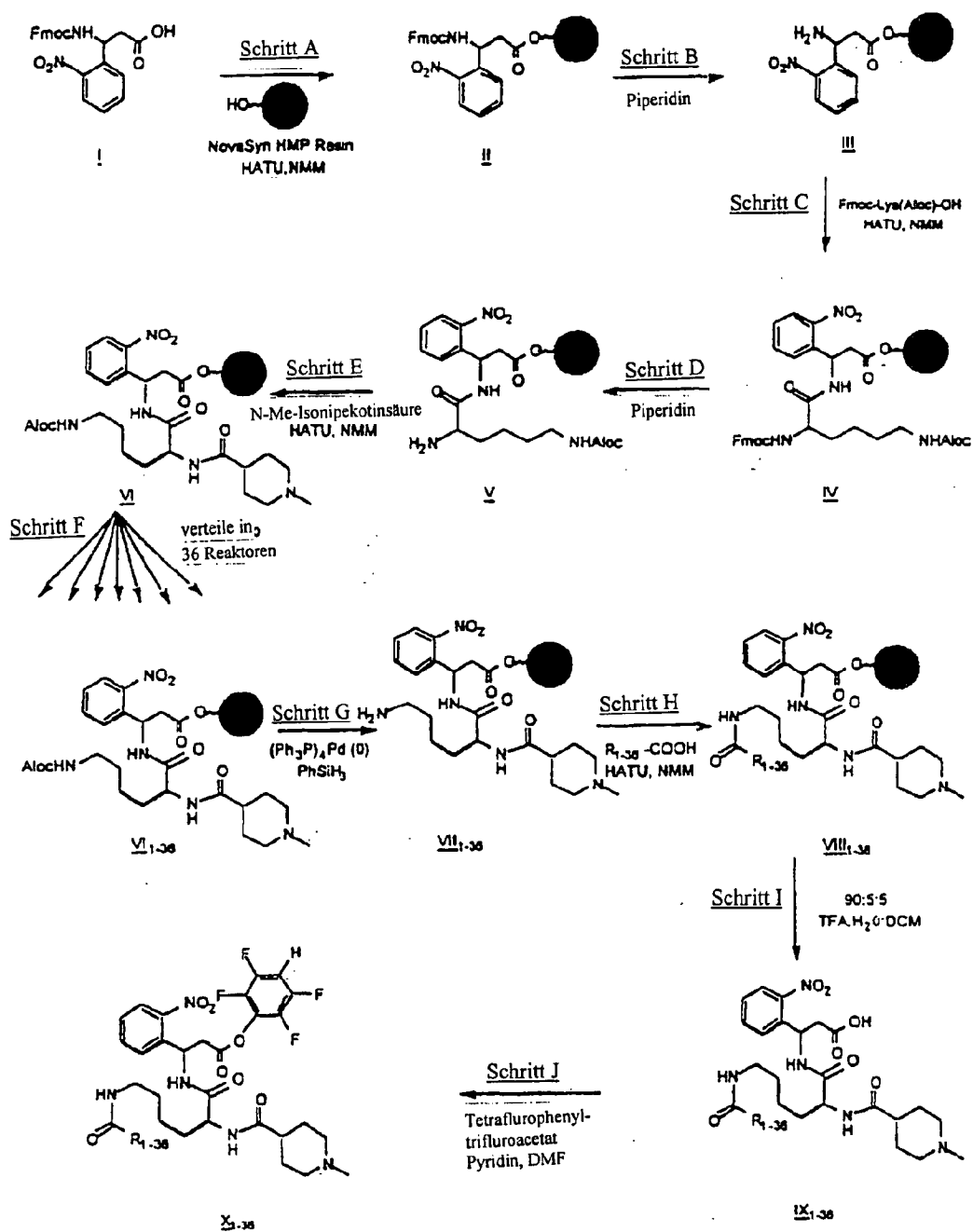




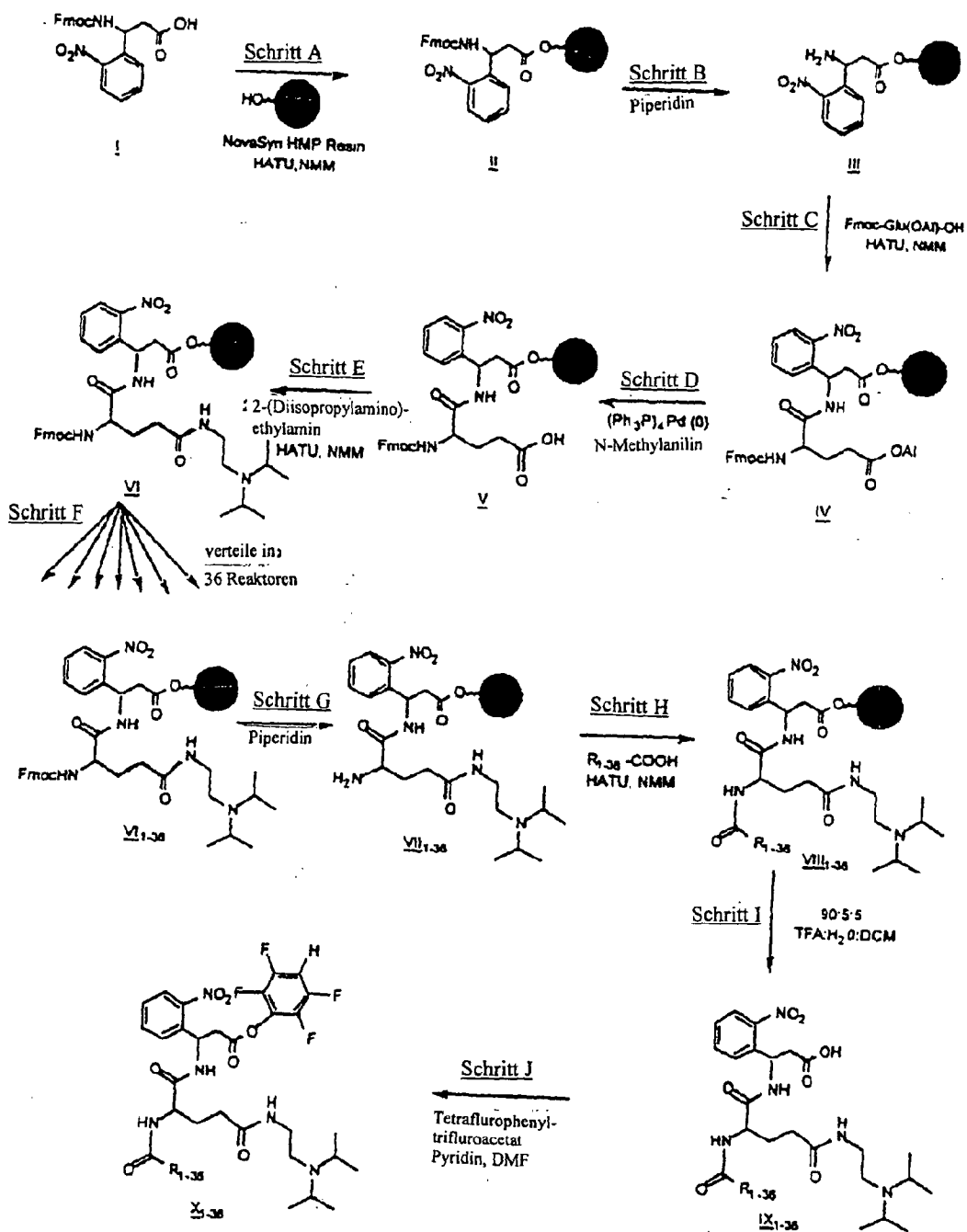
FIGUR 3



FIGUR 4

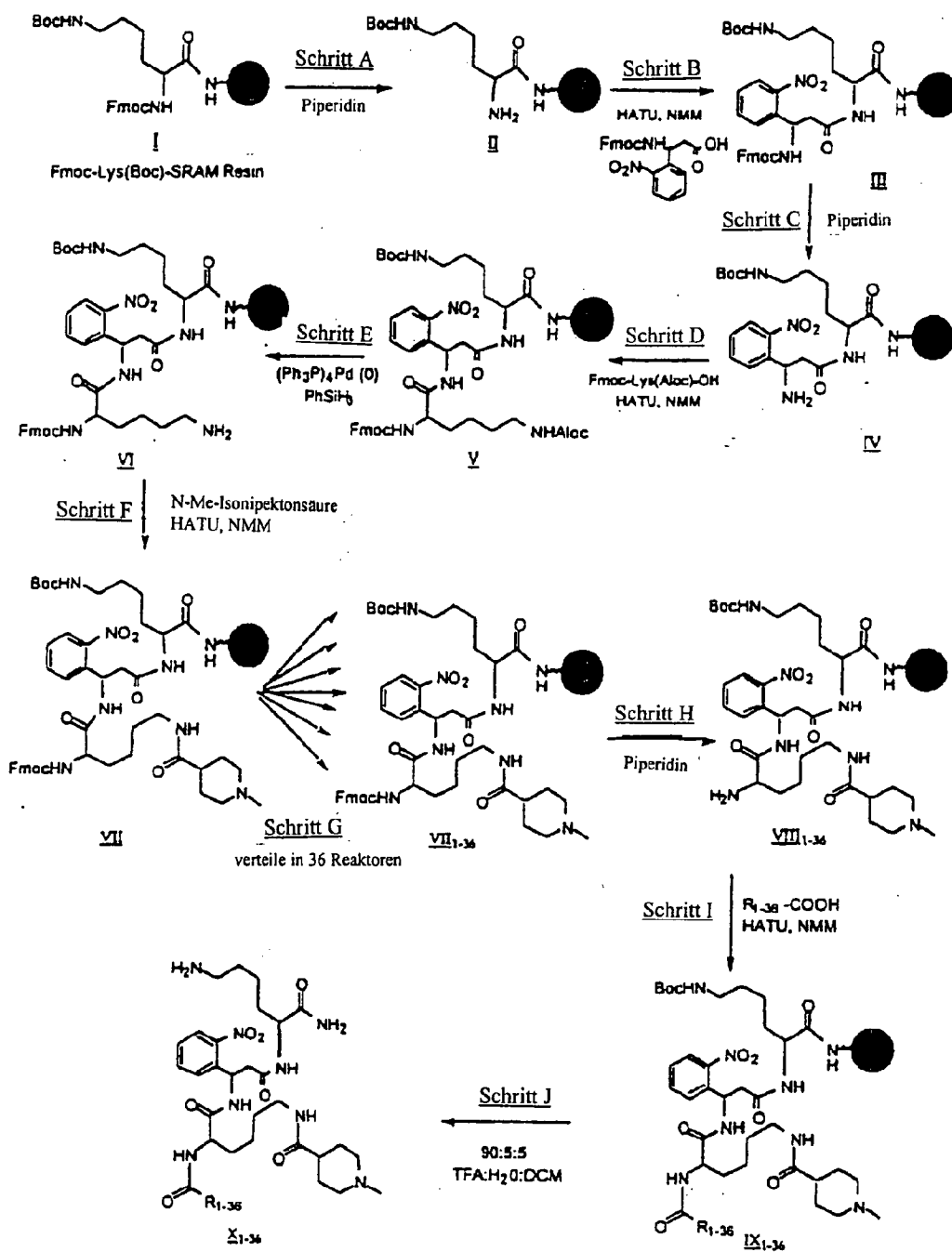


FIGUR 5

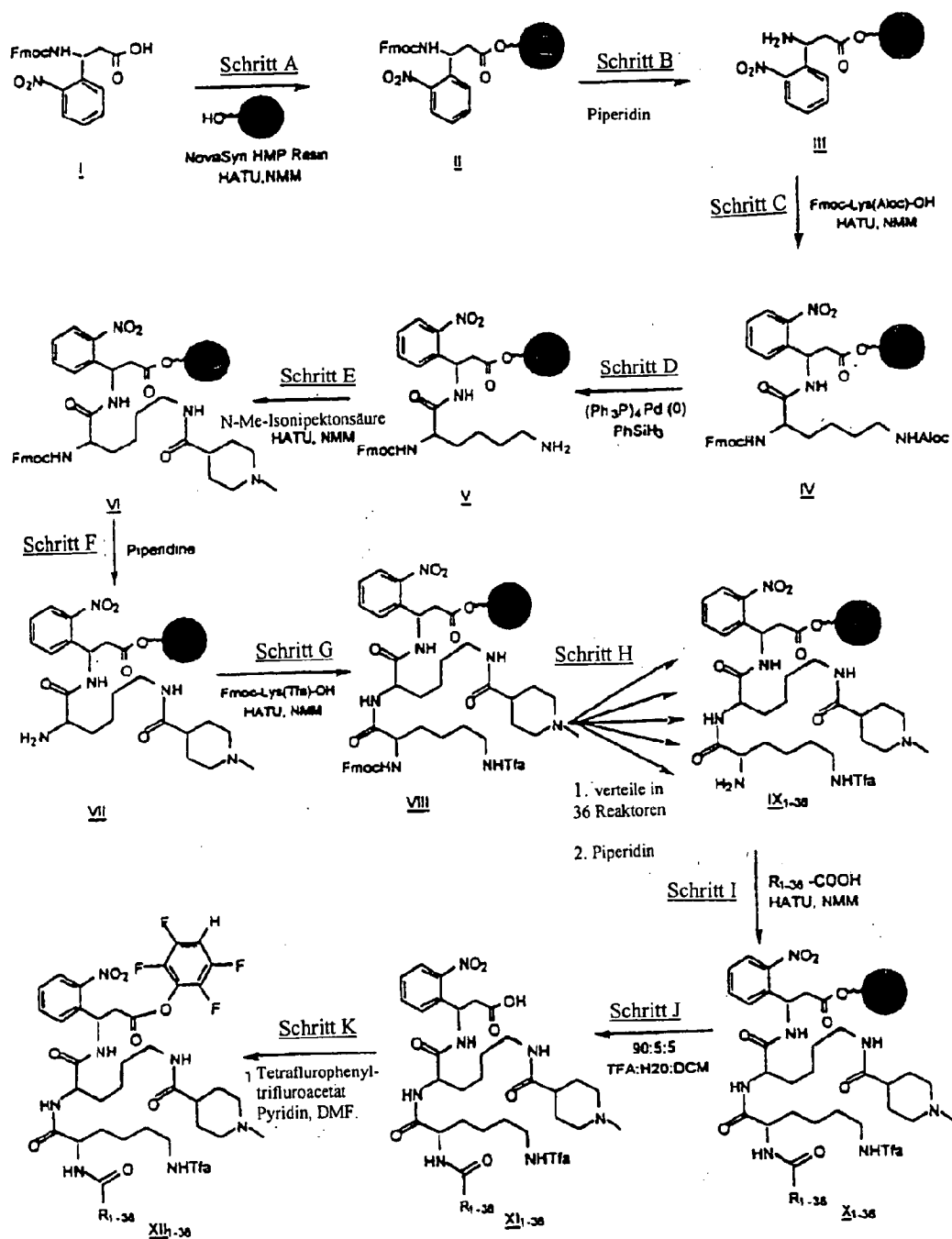


FIGUR 6

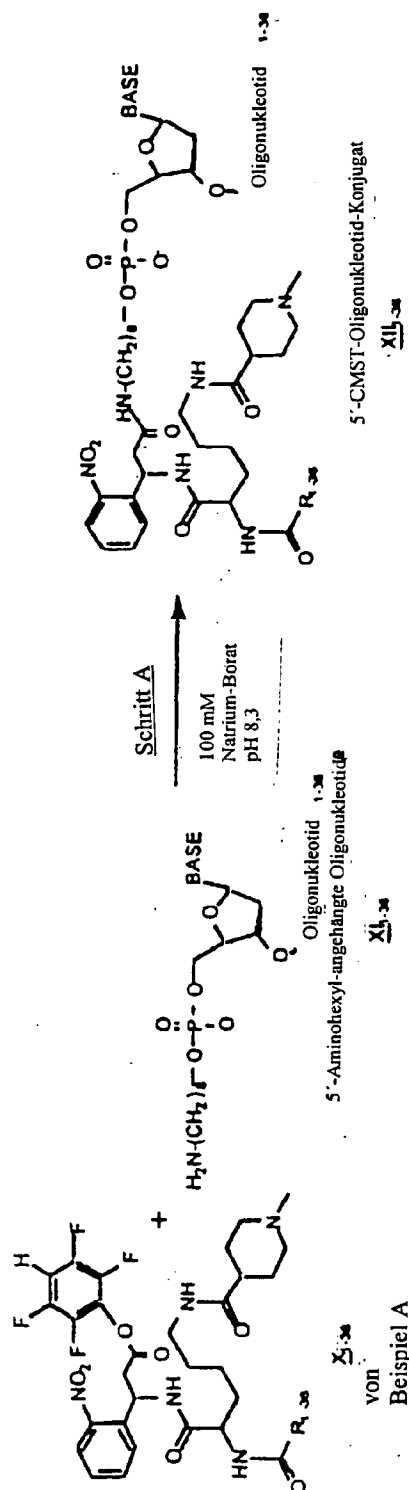




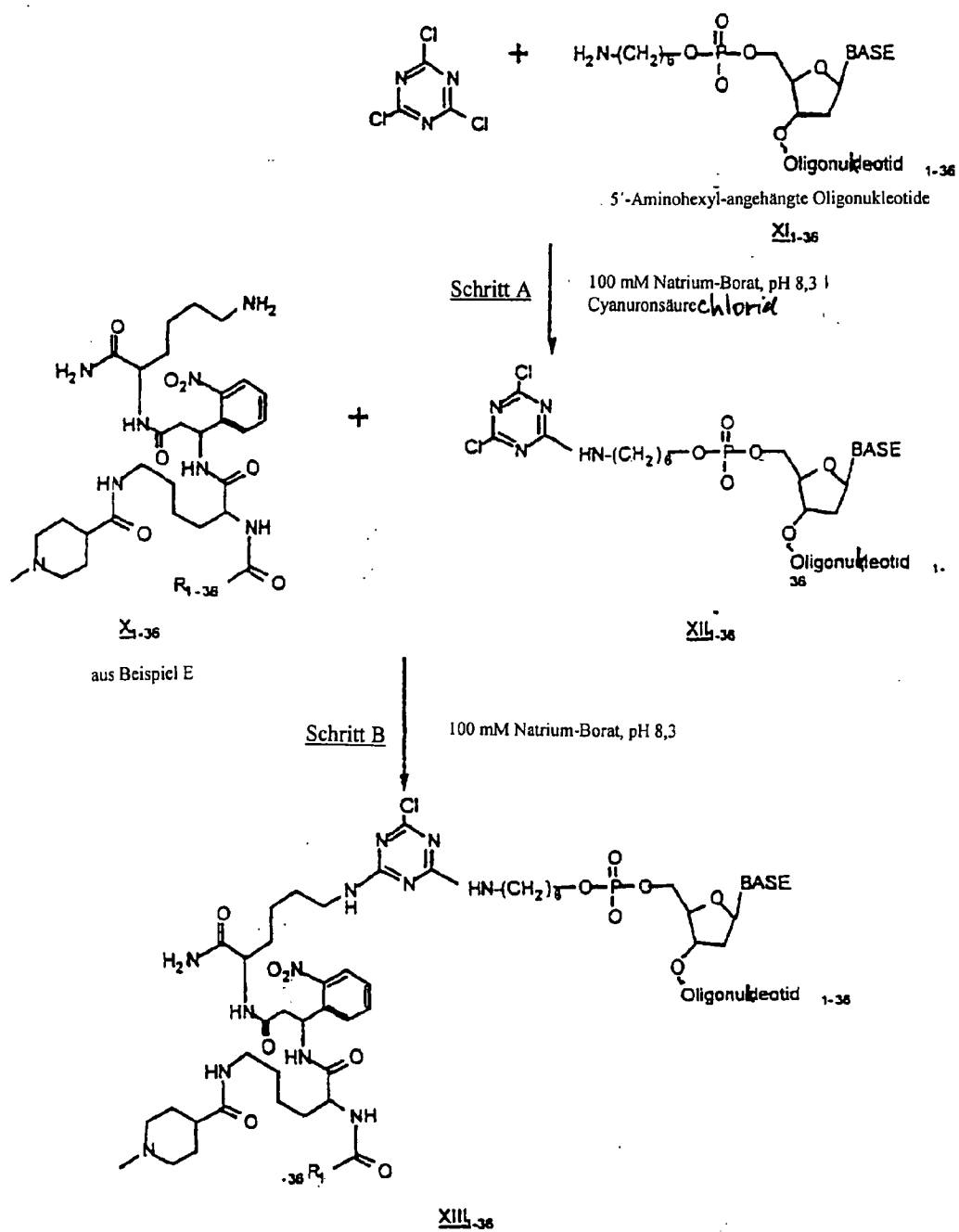
FIGUR 7



FIGUR 8

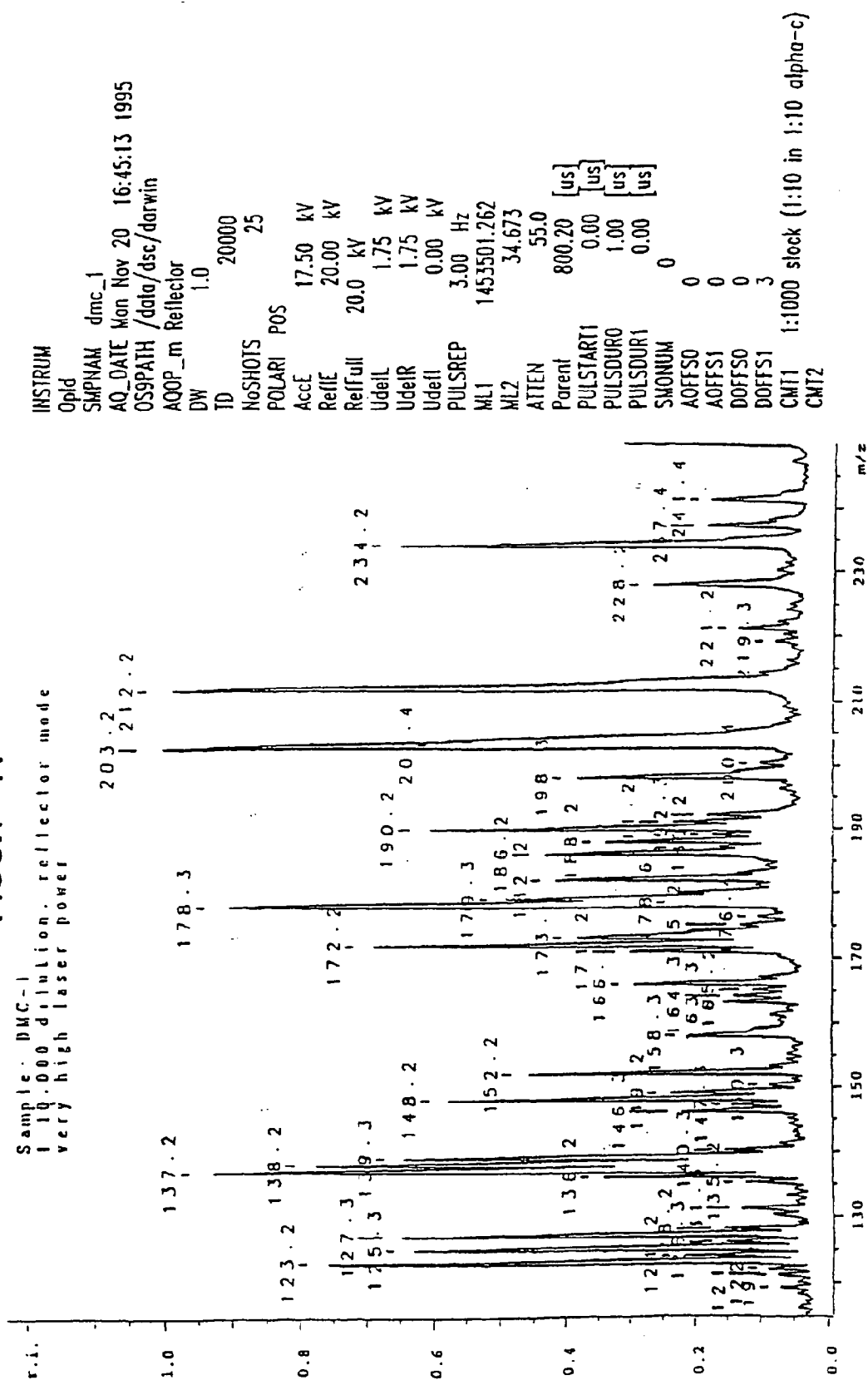


FIGUR 9

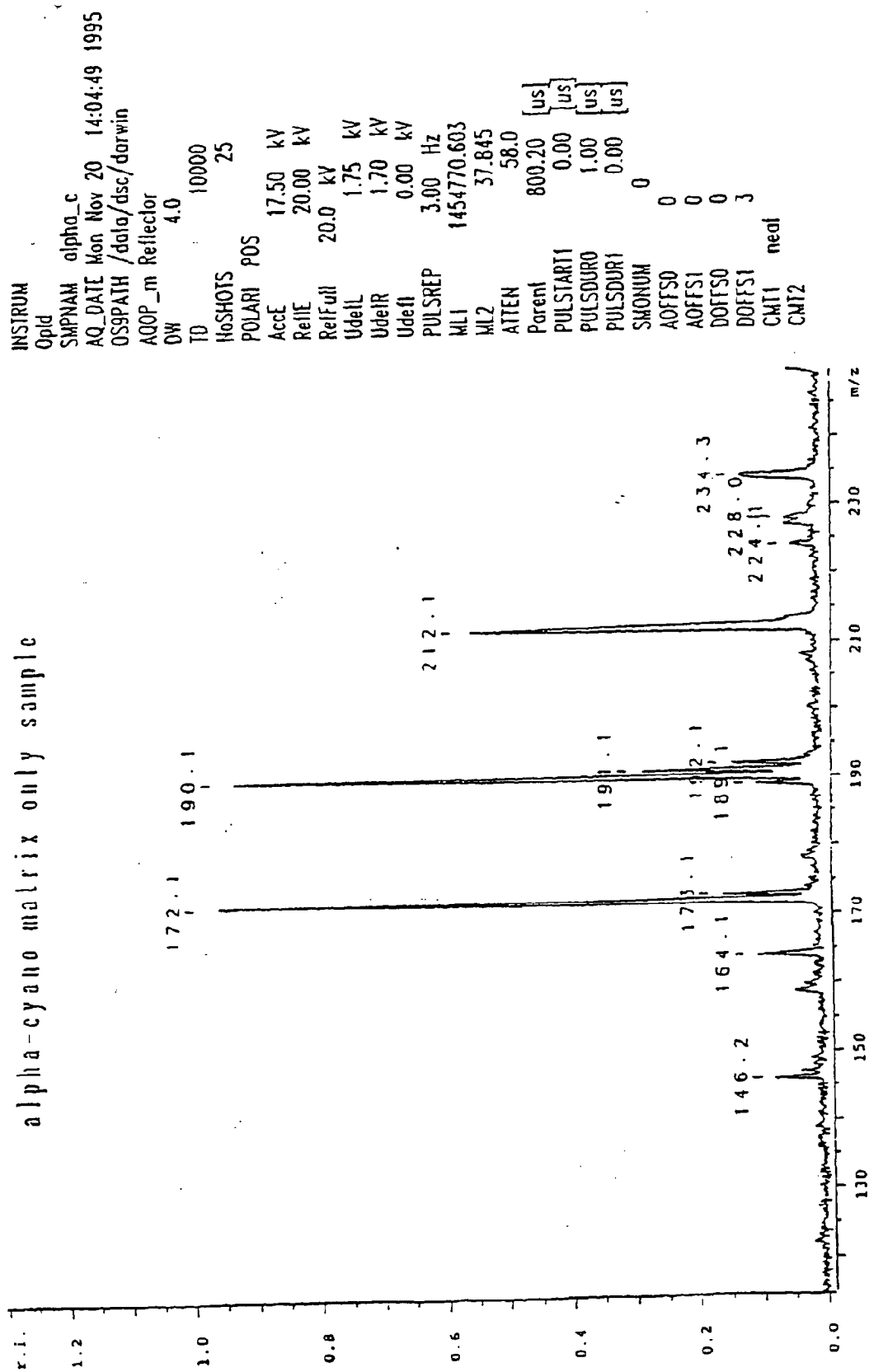


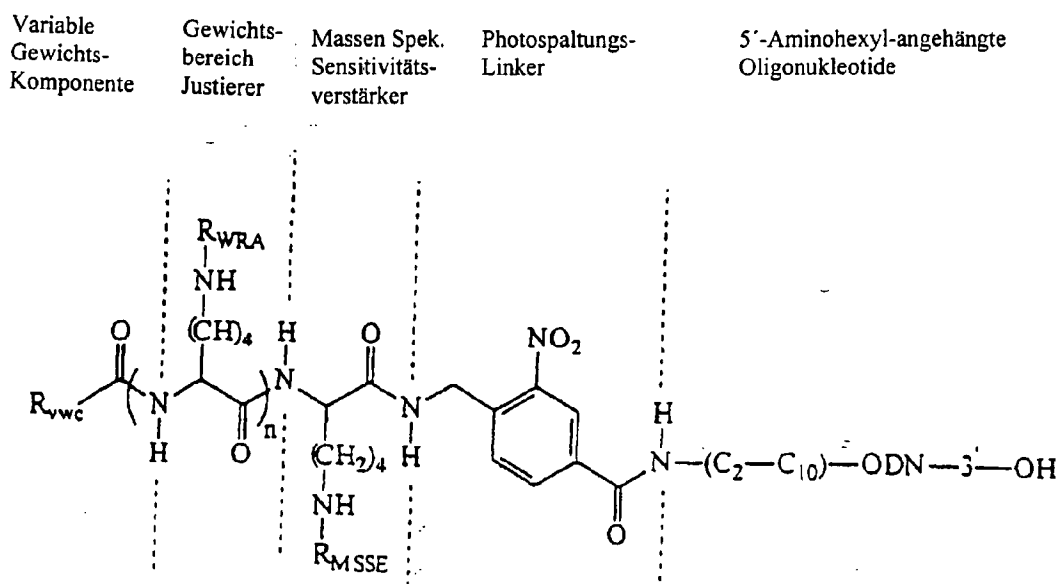
FIGUR 10

FIGUR 11

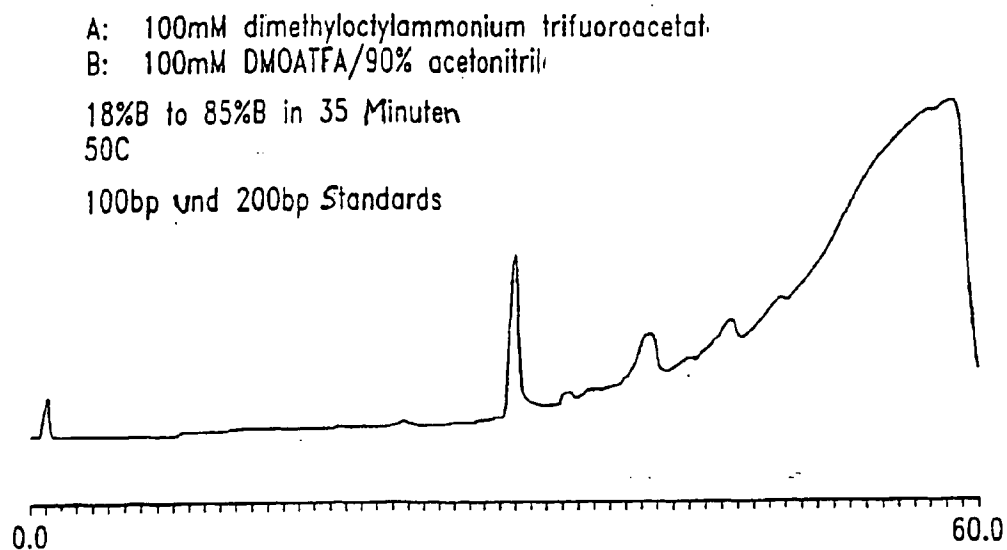


**FIGUR 12**  
alpha-cyano matrix only sample

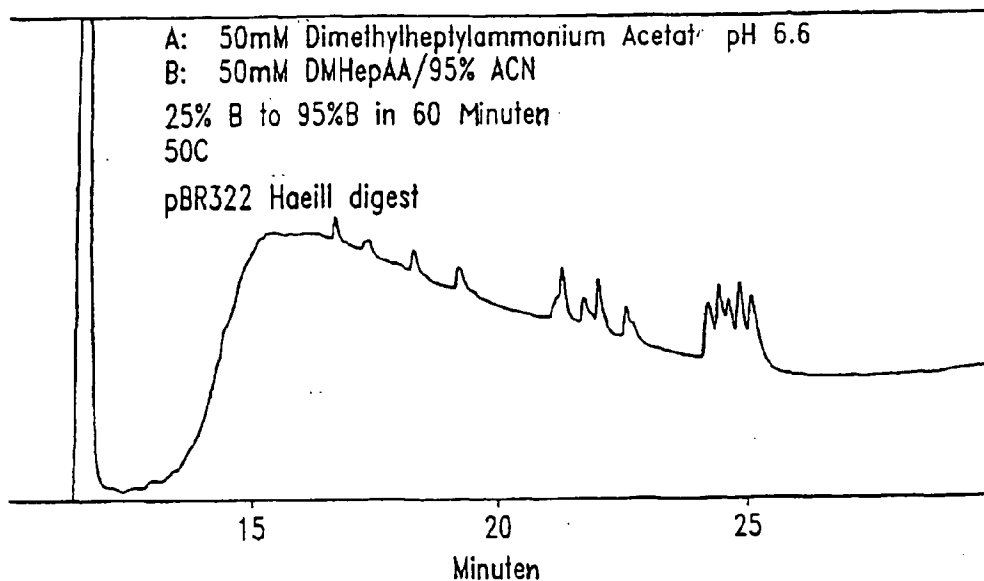




FIGUR 13

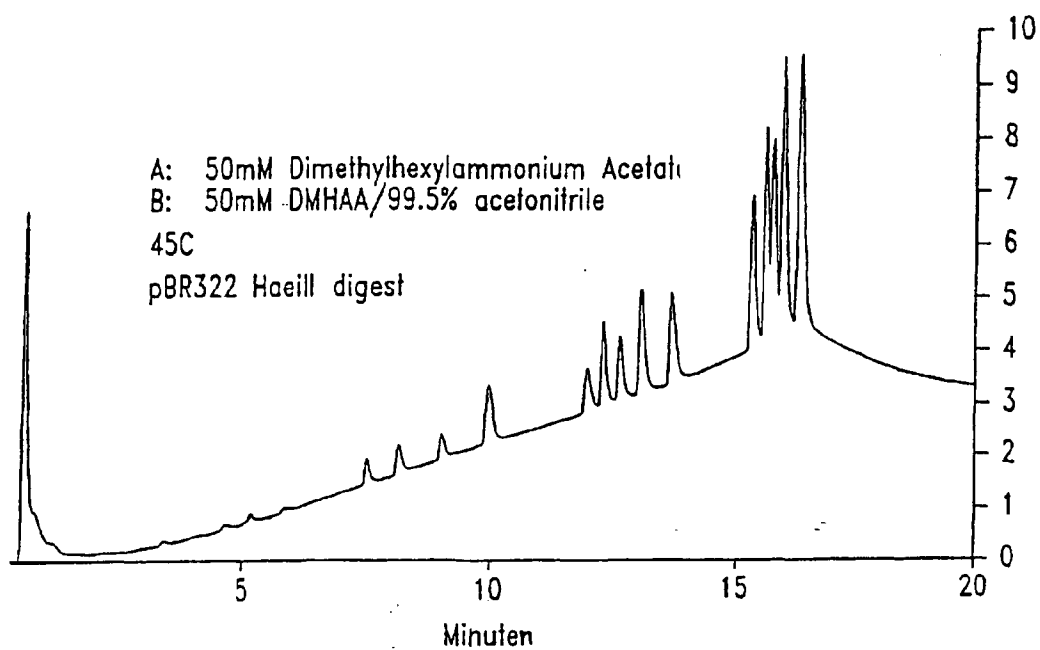


**FIGUR 14A**

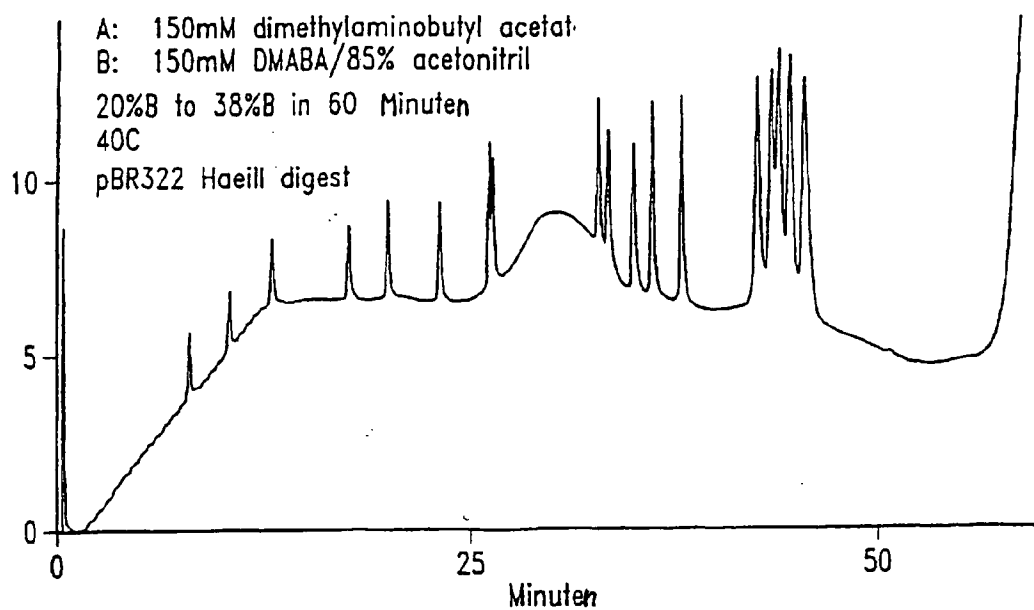


**FIGUR 14B**

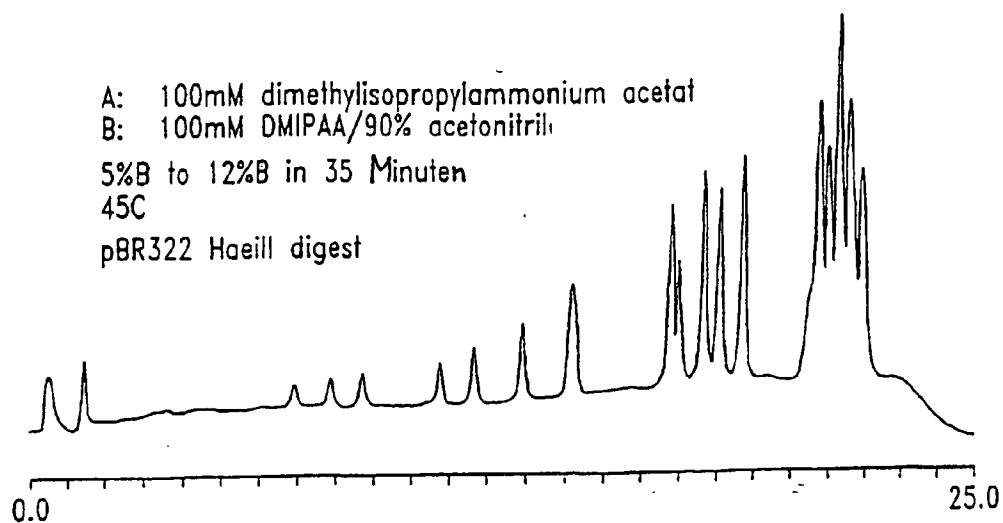




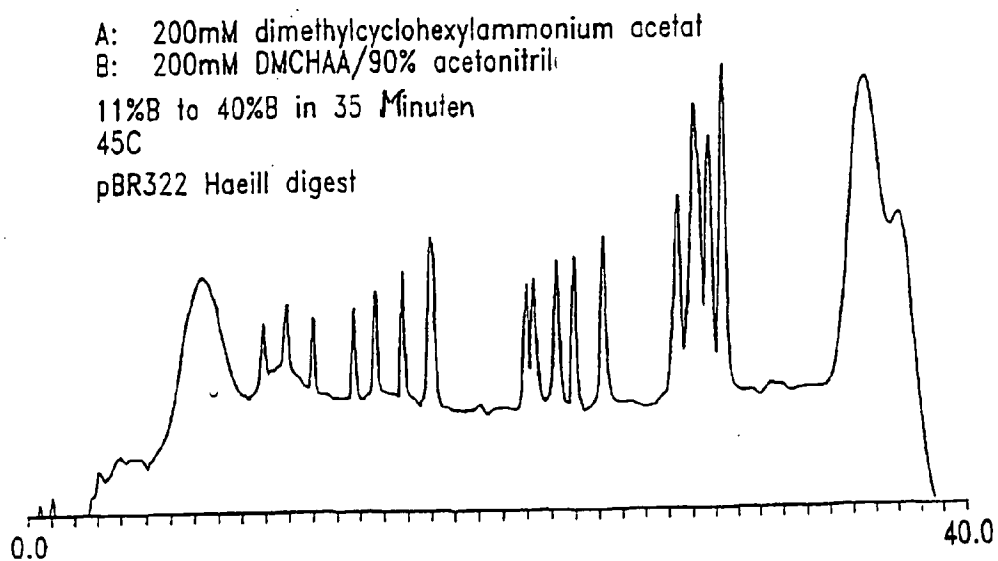
**FIGUR 14C**



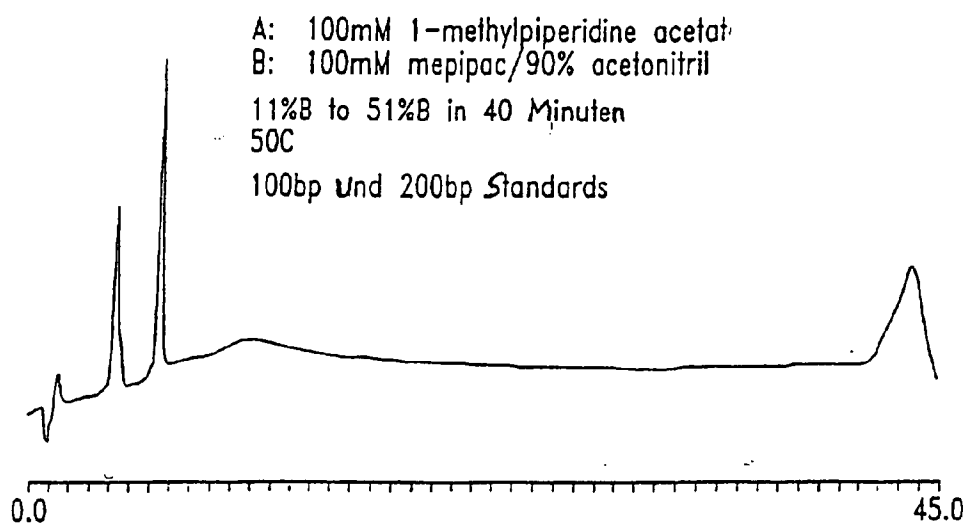
**FIGUR 14D**



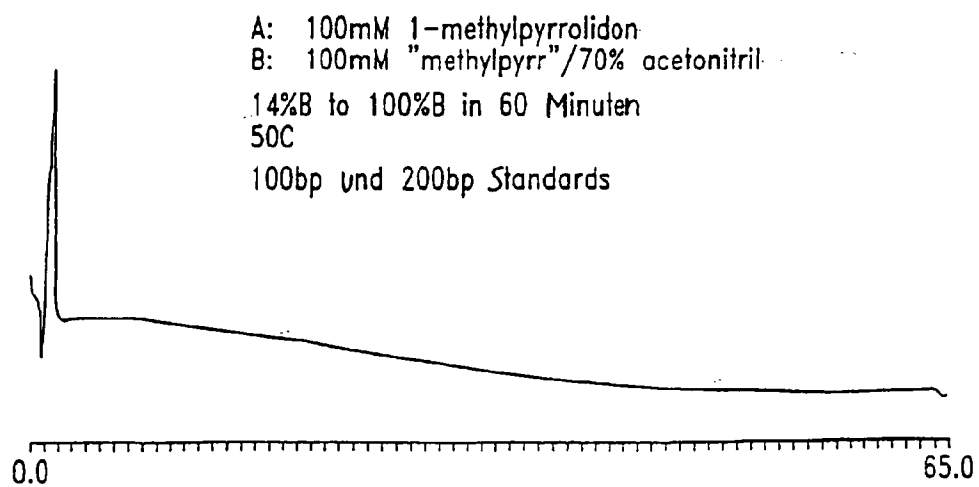
**FIGUR 14E**



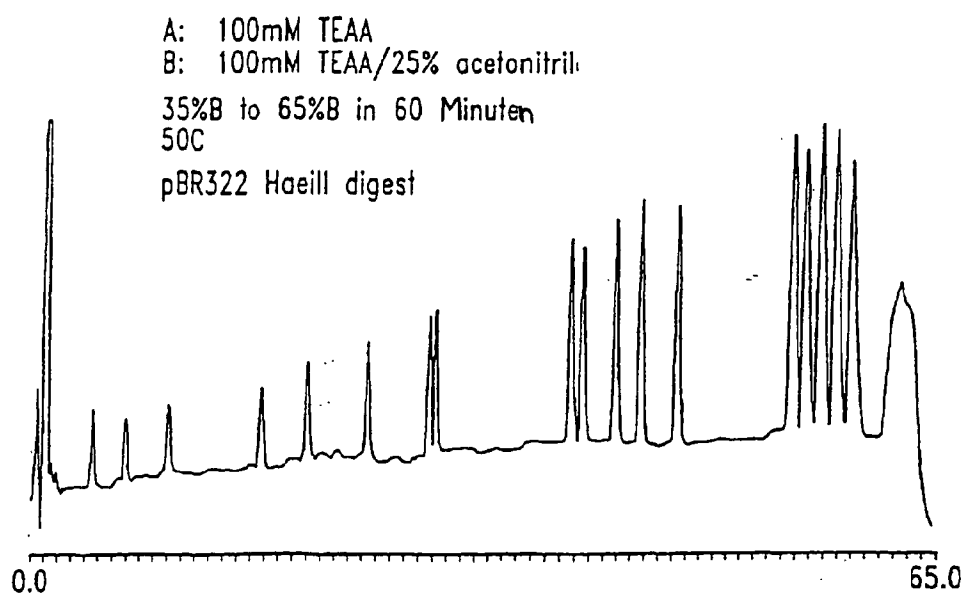
**FIGUR 14F**



**FIGUR 14G**

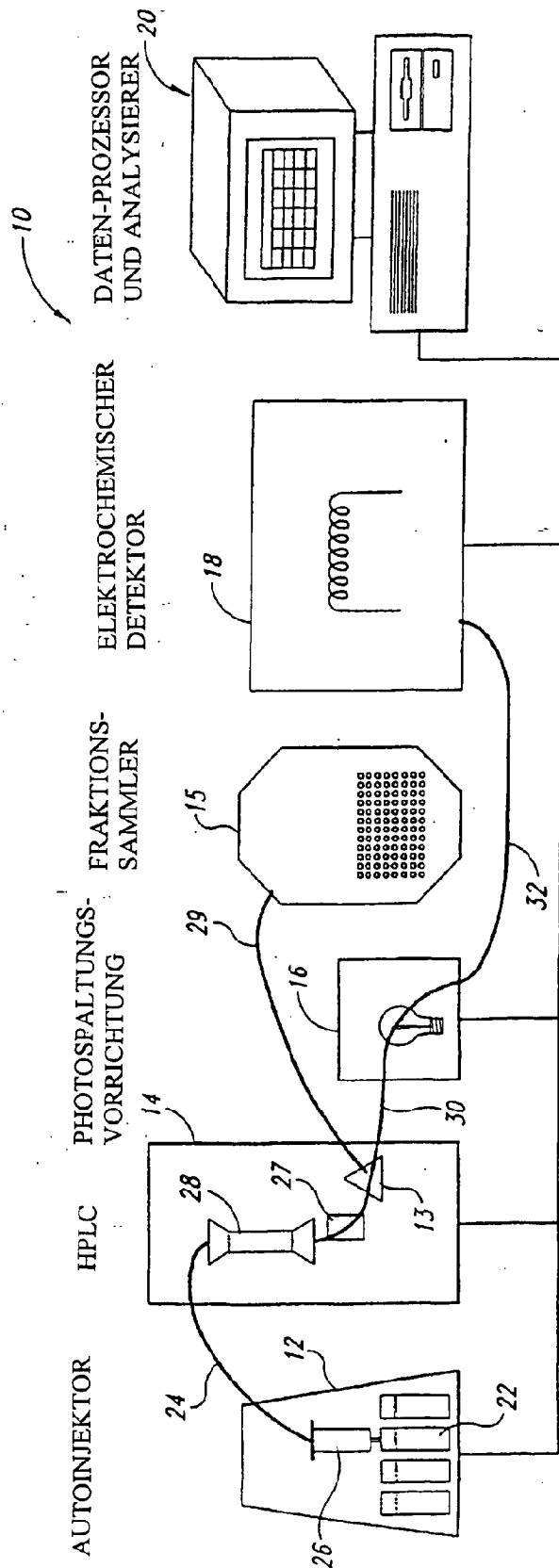


**FIGUR 14H**



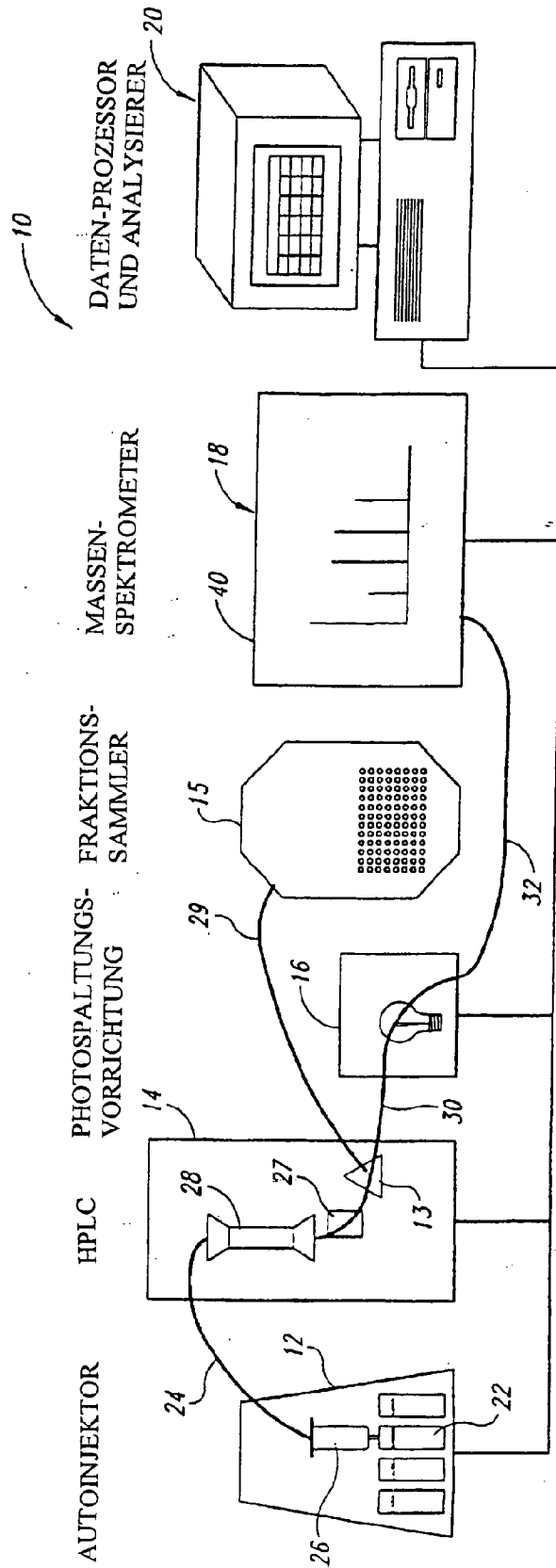
**FIGUR 14I**

GENETISCHES FINGERPRINTING UND  
DIFFERENTIAL DISPLAY SYSTEME



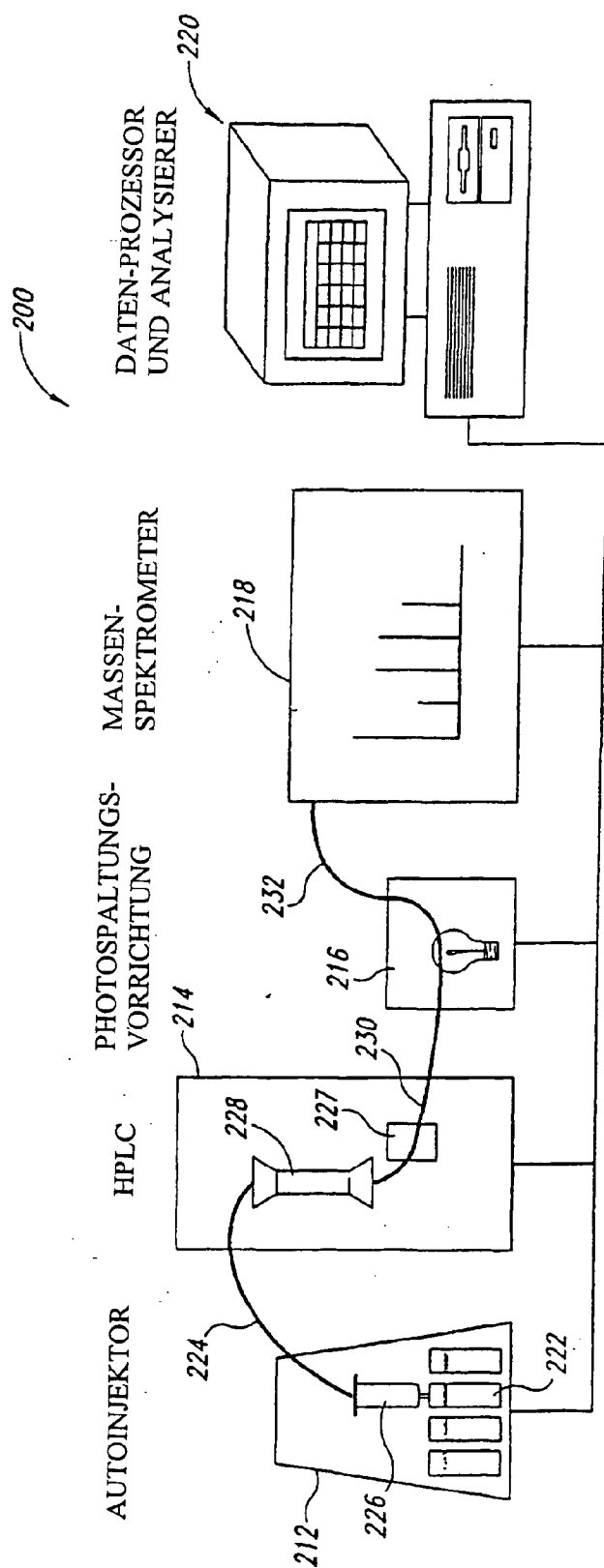
FIGUR 15

GENETISCHES FINGERPRINTING UND  
DIFFERENTIAL DISPLAY SYSTEME



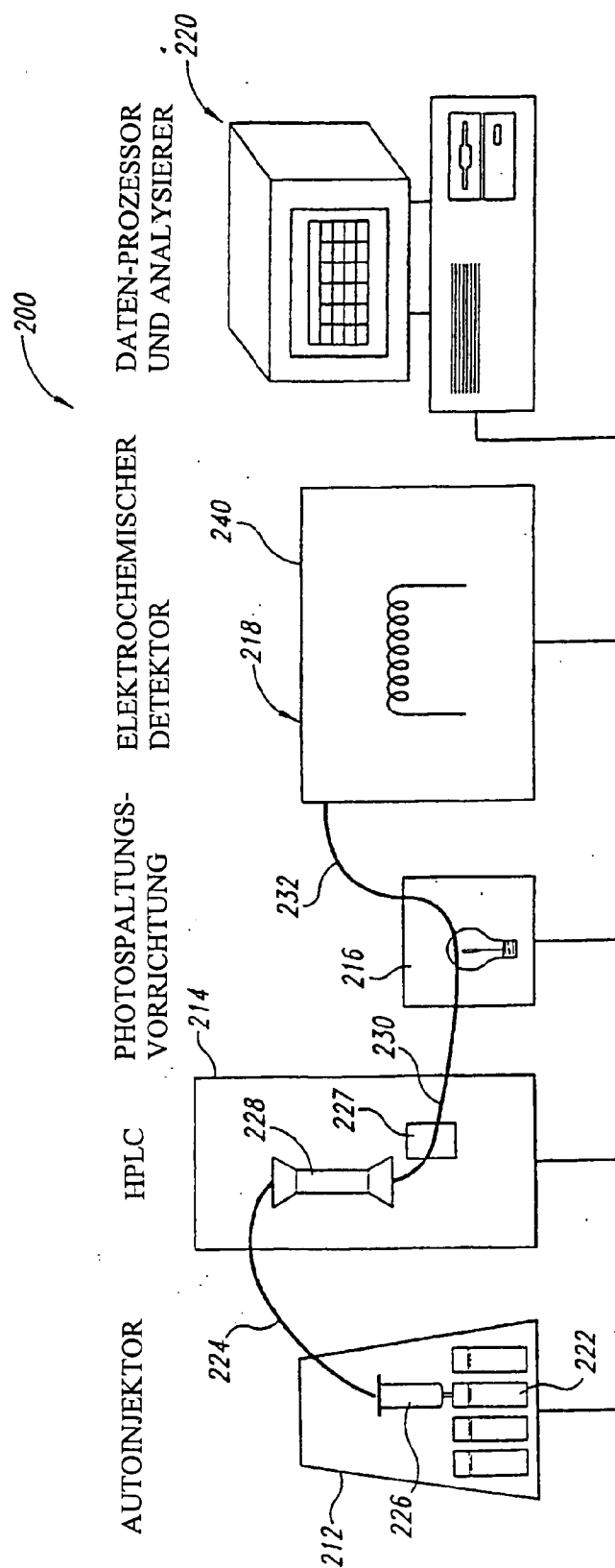
FIGUR 16

VERSUCHSSYSTEME



FIGUR 17

VERSUCHSSYSTEME



FIGUR 18



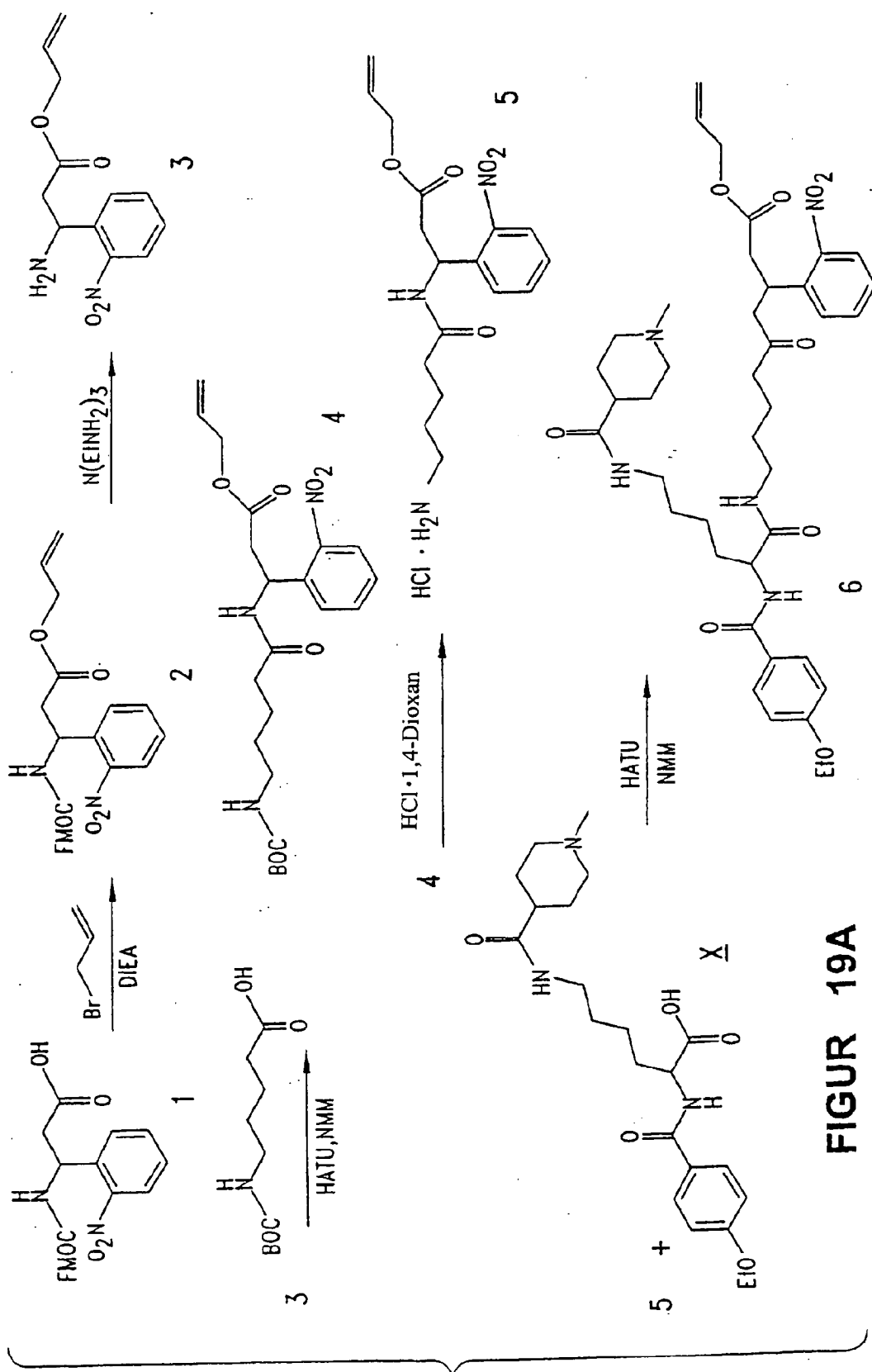
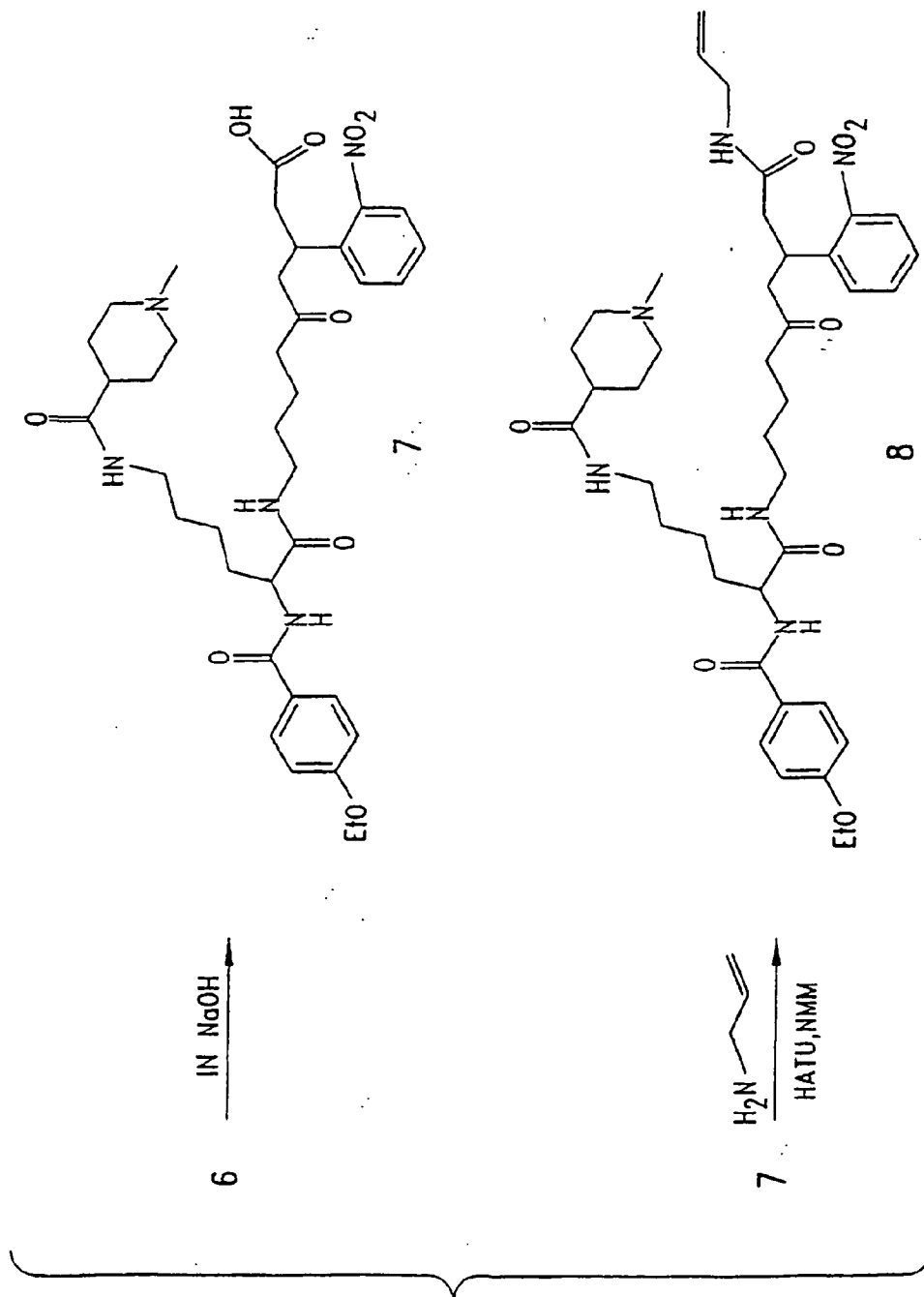
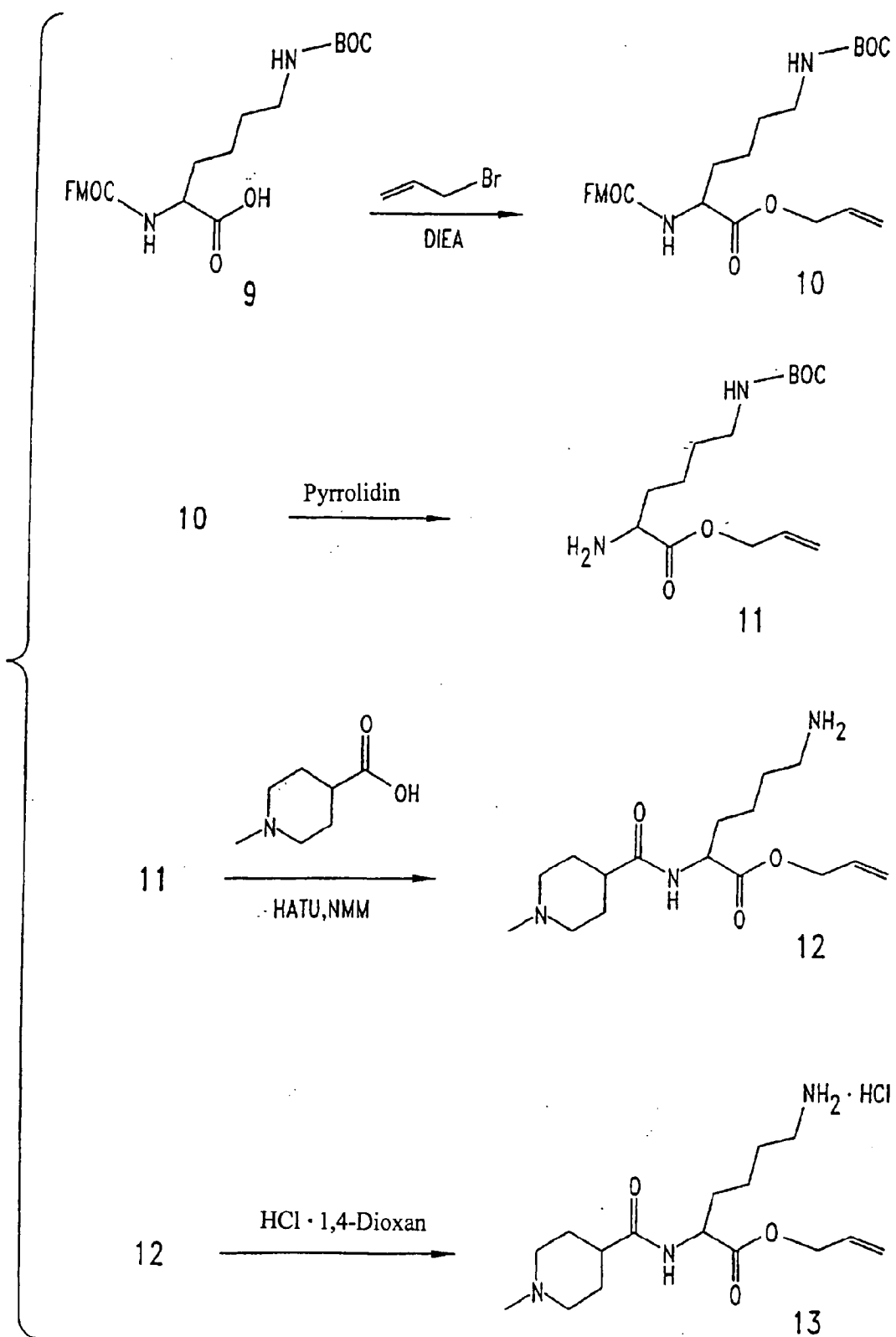


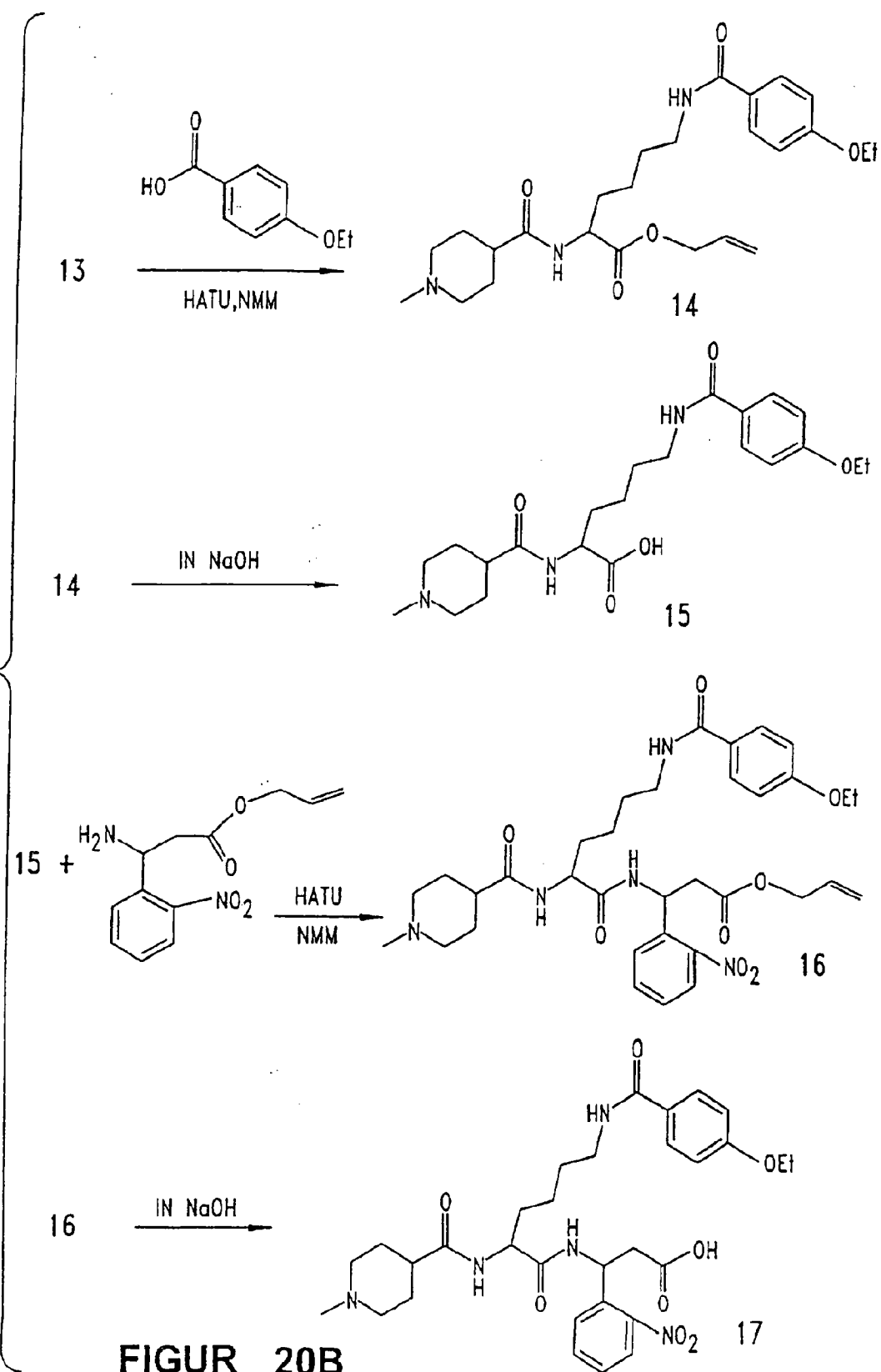
FIGURE 19A

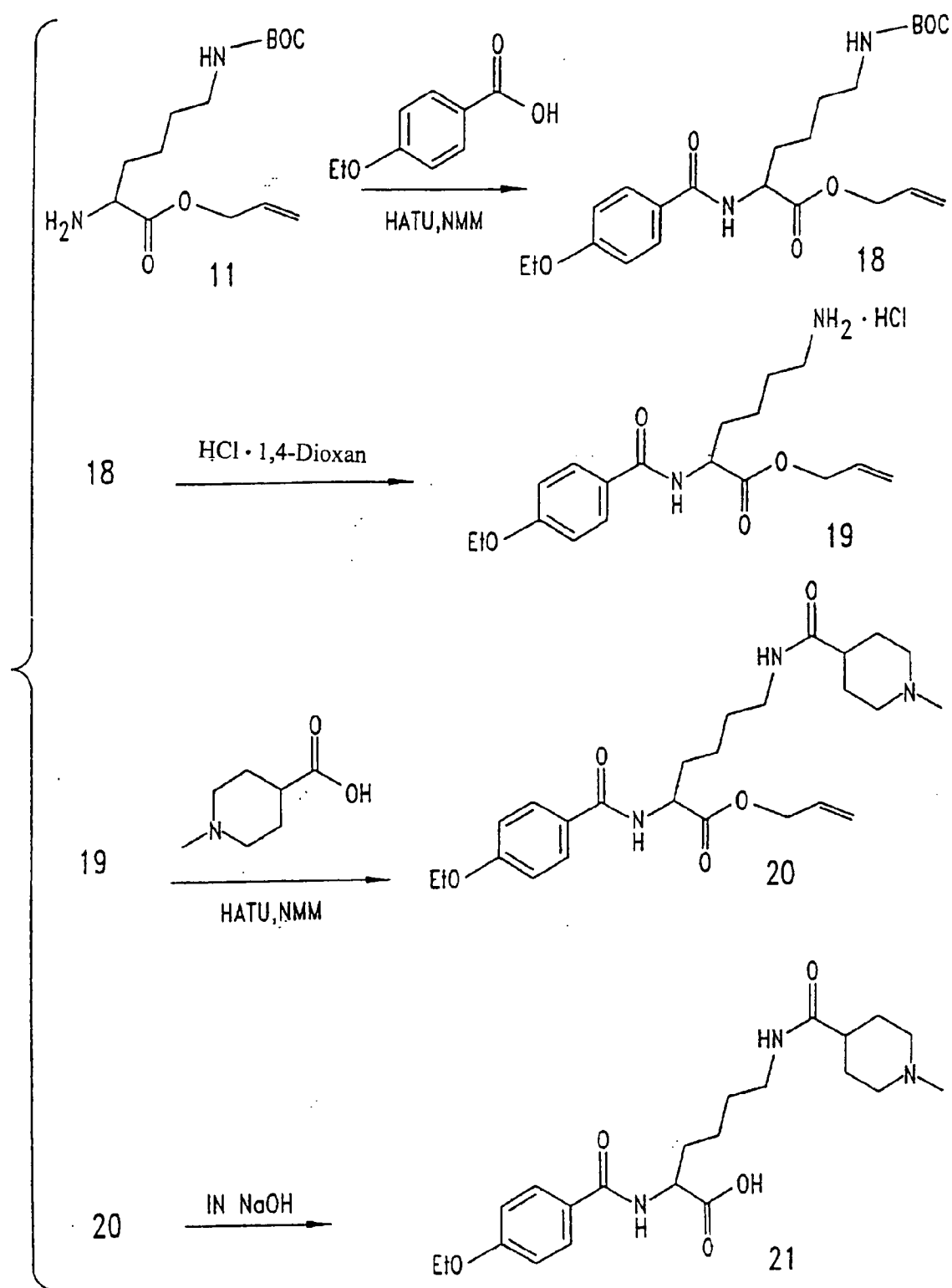


FIGUR 19B

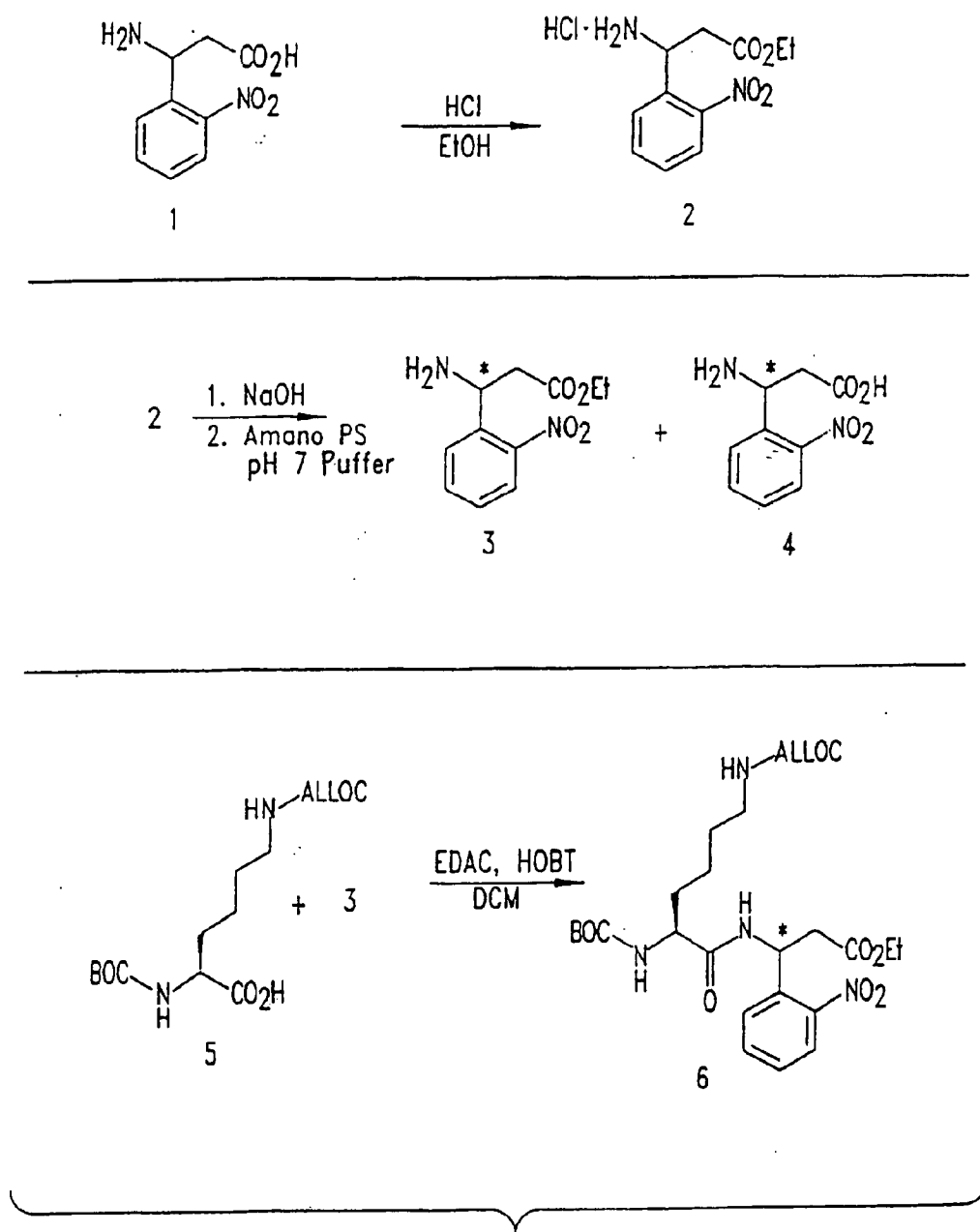


FIGUR 20A

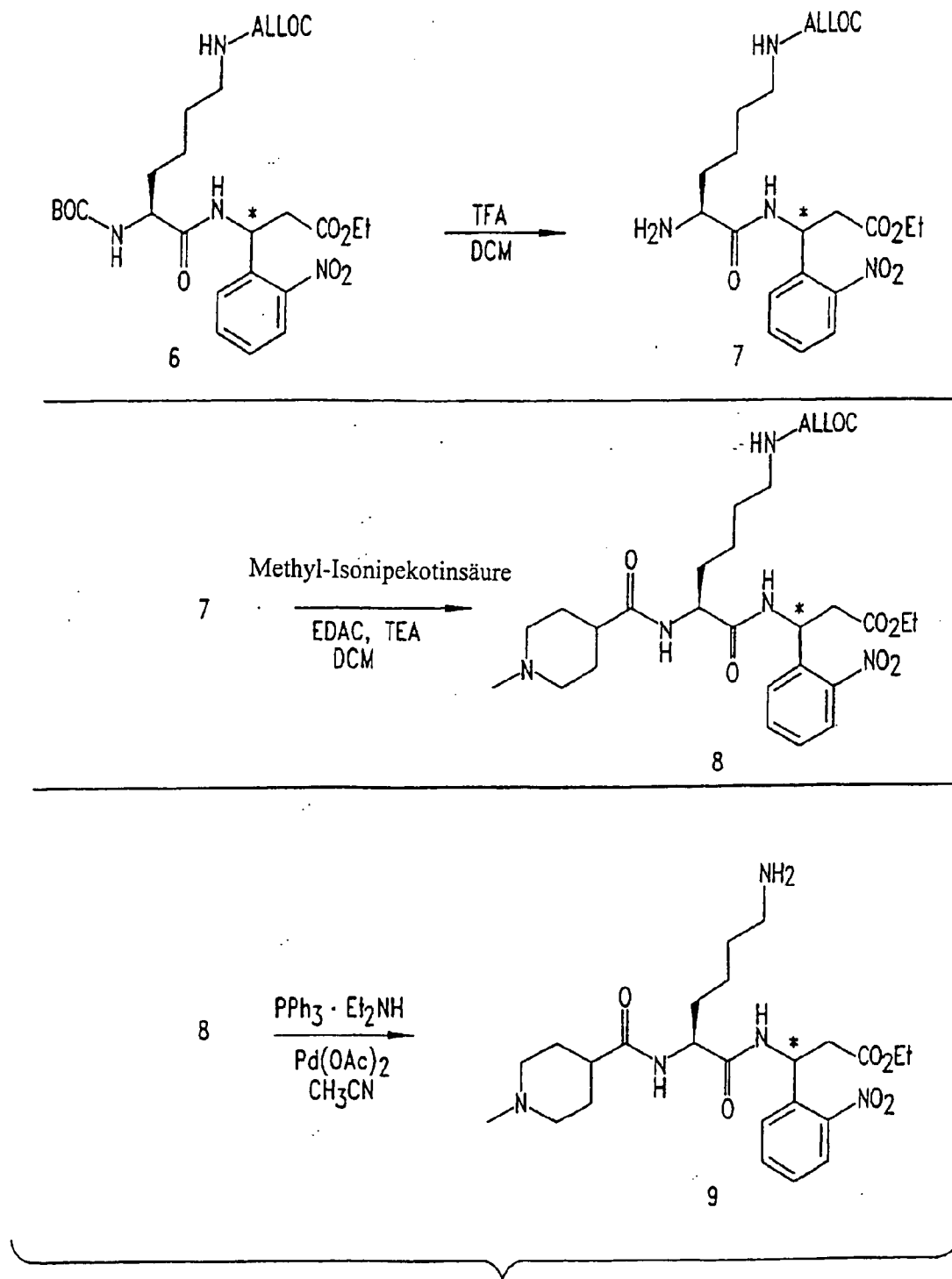




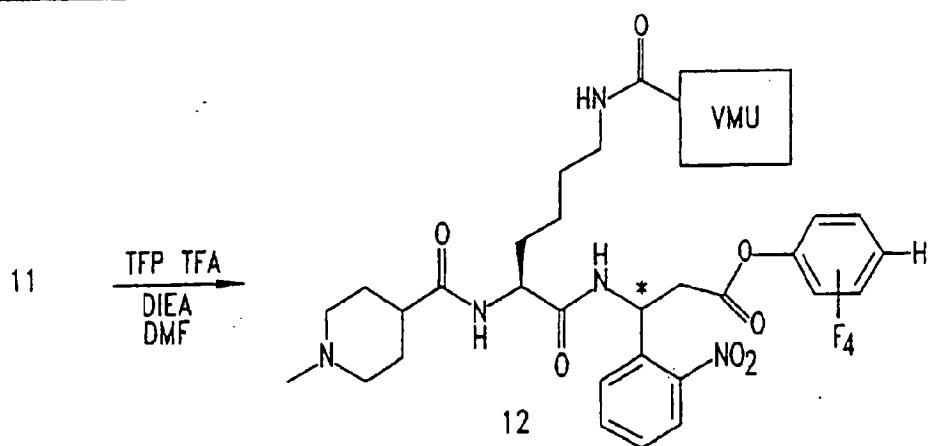
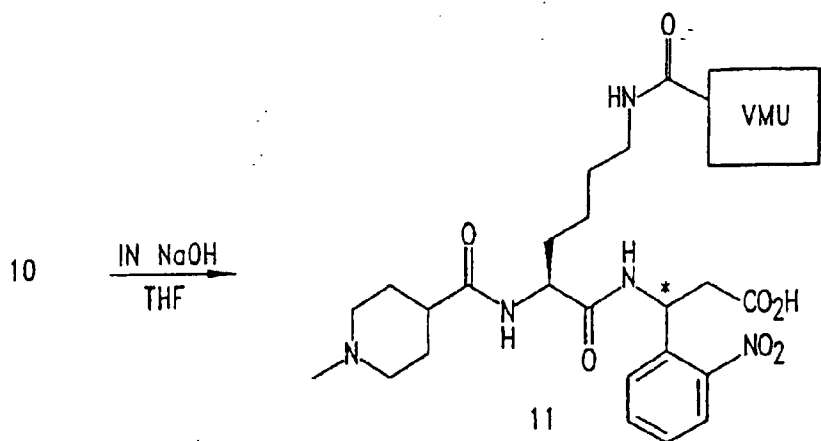
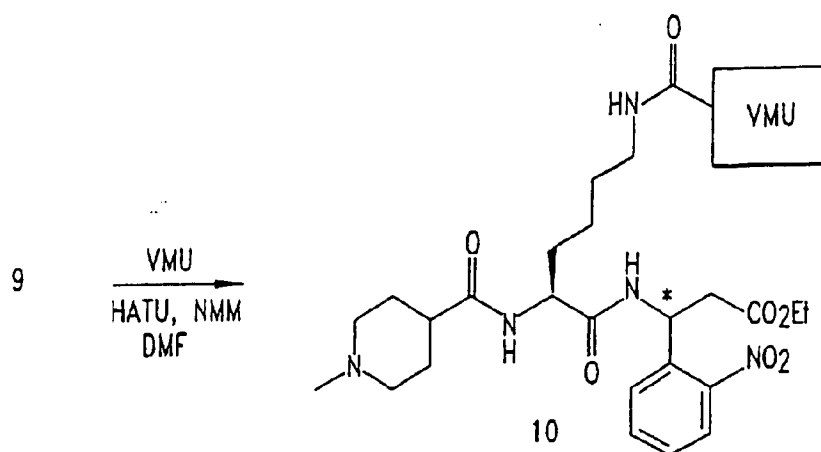
FIGUR 21



FIGUR 22A

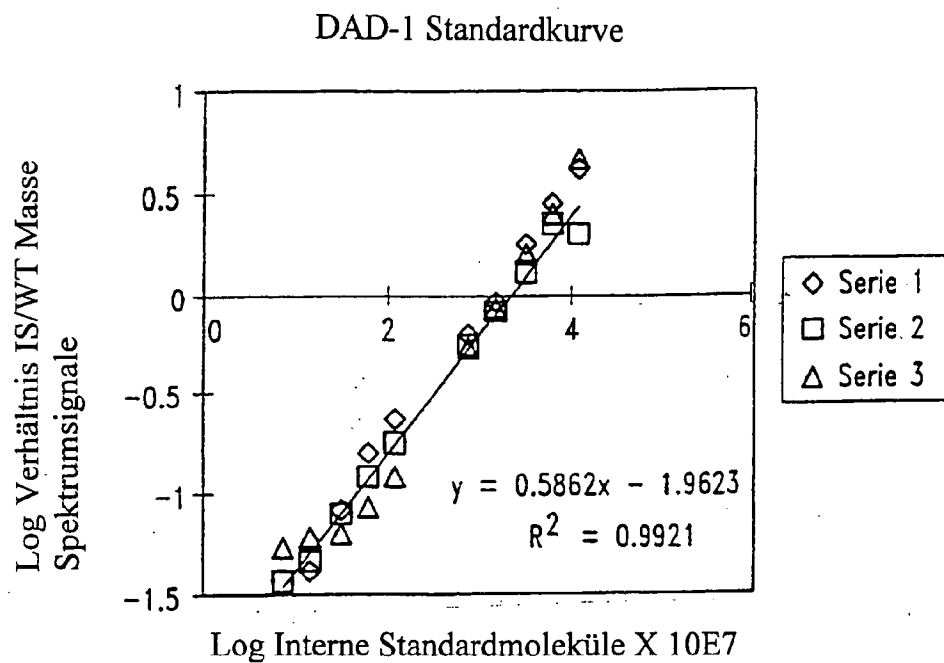


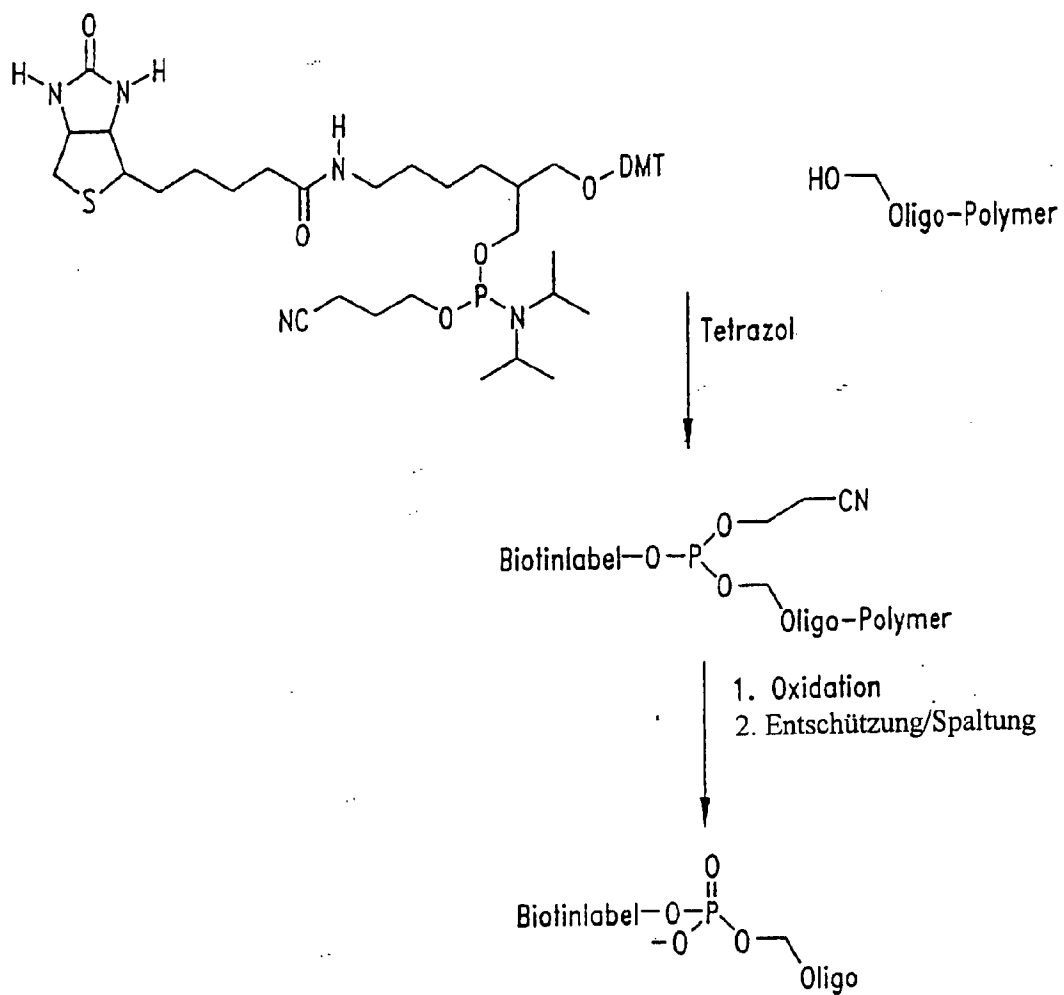
FIGUR 22B

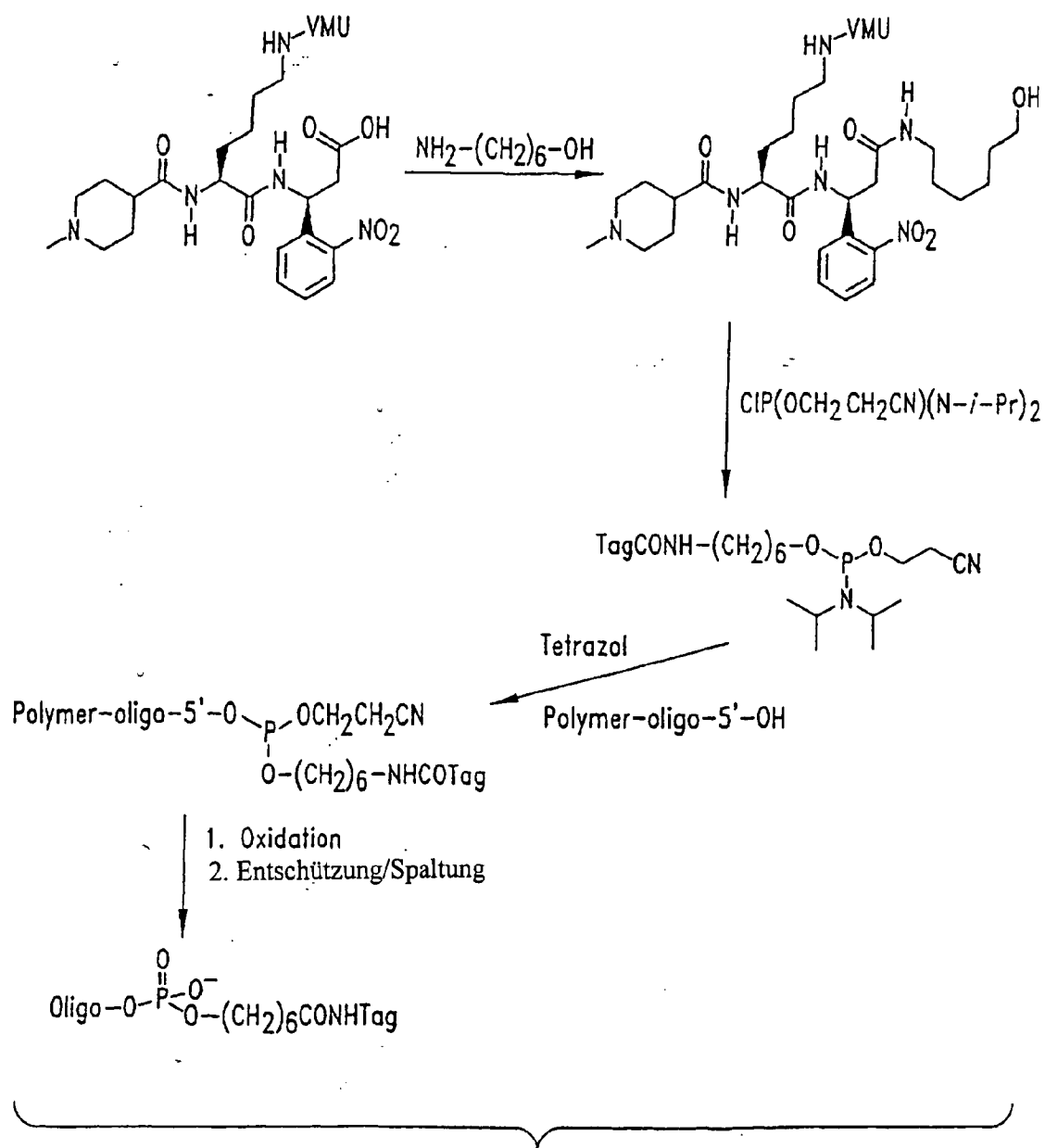


FIGUR 22C





**FIGUR 24**



FIGUR 25

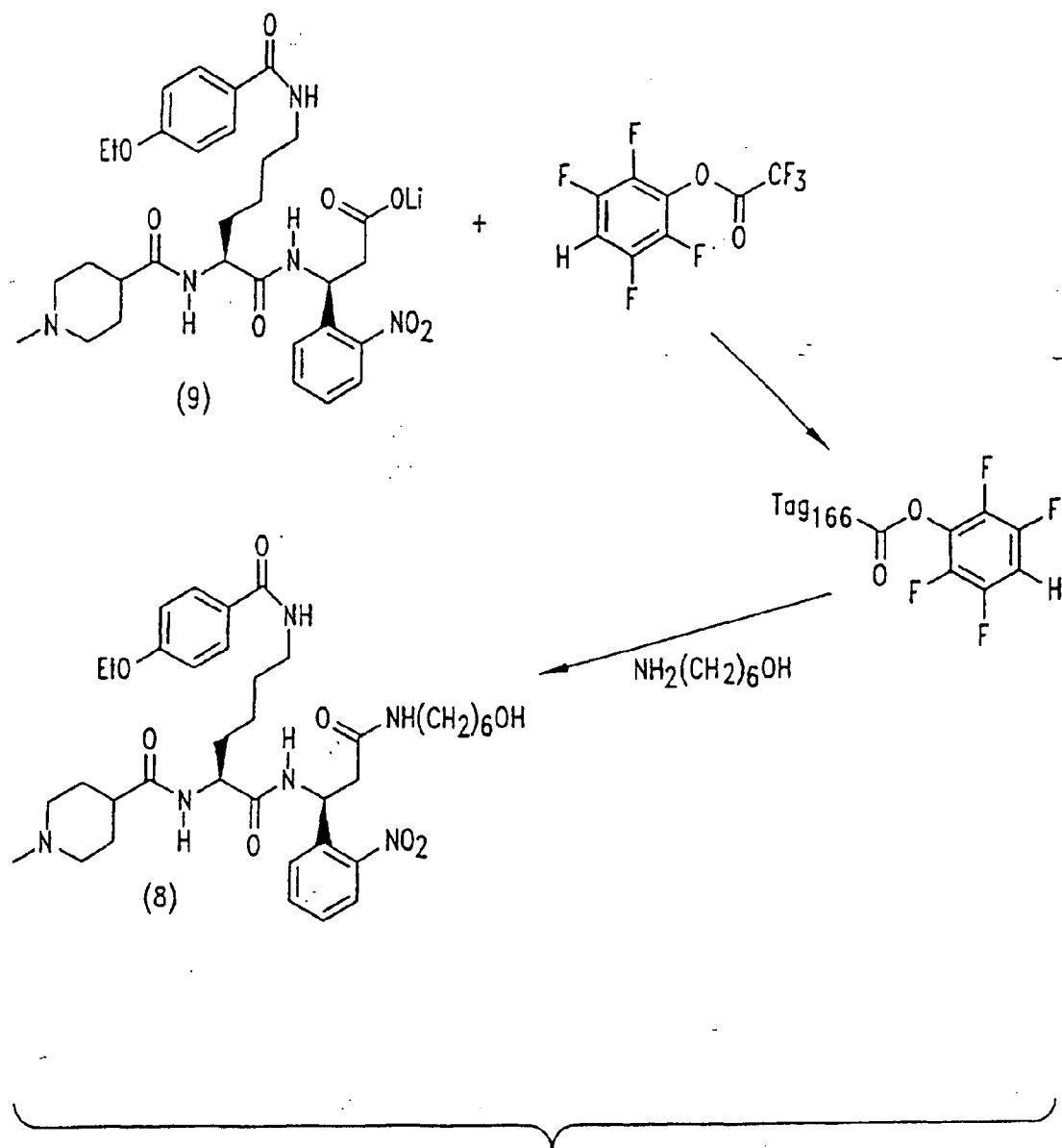
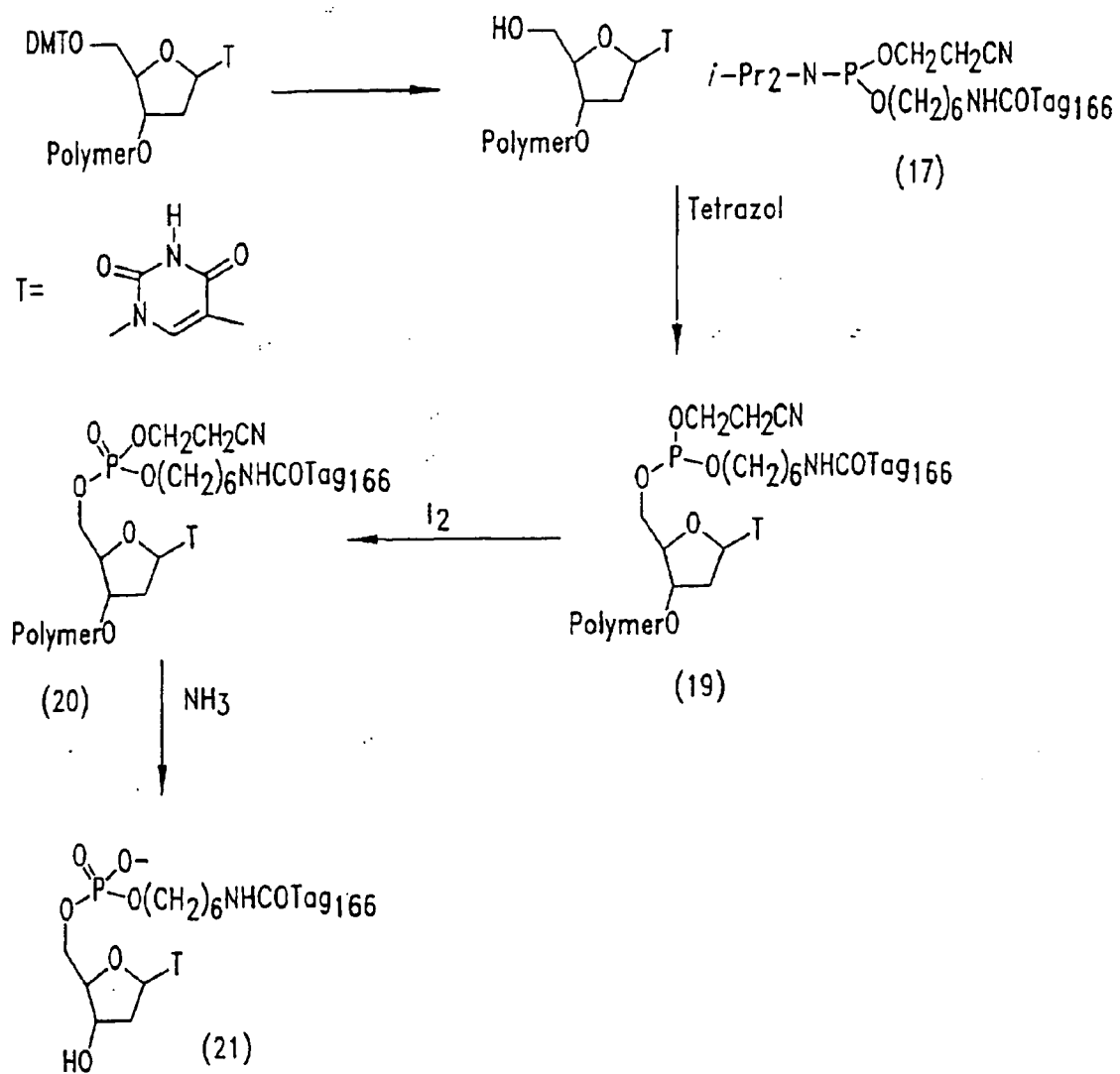
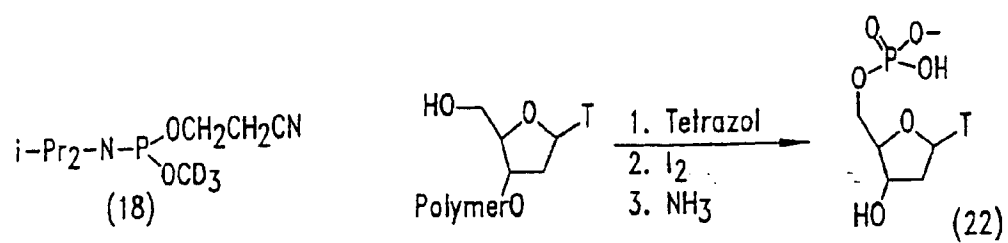


FIGURE 26



FIGUR 27



**FIGUR 28**



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : <b>C07H 21/00, C12Q 1/68</b></p>	<p><b>A3</b></p>	<p>(11) International Publication Number: <b>WO 99/05319</b> (43) International Publication Date: 4 February 1999 (04.02.99)</p>
<p>(21) International Application Number: PCT/US98/15008 (22) International Filing Date: 22 July 1998 (22.07.98) (30) Priority Data: 08/898,180 22 July 1997 (22.07.97) US 08/898,564 22 July 1997 (22.07.97) US 08/898,501 22 July 1997 (22.07.97) US (71) Applicant (for all designated States except US): RAPIGENE, INC. [US/US]; 1631 - 220th Street Southeast, Bothell, WA 98021 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): VAN NESS, Jeffrey [US/US]; 10020 - 49th Avenue Northeast, Seattle, WA 98125 (US). TABONE, John, C. [US/US]; 12117 Northeast 166th Place, Bothell, WA 98011 (US). HOWBERT, Jeffrey [US/US]; 12740 Northeast 30th Street, Bellevue, WA 98005 (US). MULLIGAN, John, T. [US/US]; 5823 - 17th Avenue Northeast, Seattle, WA 98105 (US). (74) Agents: PARKER, David, W. et al.; Seed and Berry LLP, 6300 Columbia Center, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).</p>		<p>(81) Designated States: AL, AM, AT, AU, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> (88) Date of publication of the international search report: 14 May 1999 (14.05.99)</p>
<p>(54) Title: METHODS AND COMPOUNDS FOR ANALYZING NUCLEIC ACIDS BY MASS SPECTROMETRY</p>		
<p>(57) Abstract  Tags and linkers specifically designed for a wide variety of nucleic acid reactions are disclosed, which are suitable for a wide variety of nucleic acid reactions wherein separation of nucleic acid molecules based upon size is required.</p>		

evaporated *in vacuo*. Compound III is isolated by flash chromatography ( $\text{CH}_2\text{Cl}_2 \rightarrow \text{EtOAc}$ ).

Step C. Compound III is dissolved in  $\text{CH}_2\text{Cl}_2$ ,  $\text{Pd}(\text{PPh}_3)_4$  (0.07 eq.) and N-methylaniline (2 eq.) are added, and the mixture stirred at room temperature for 4 hr. The mixture is diluted with  $\text{CH}_2\text{Cl}_2$ , washed with 1 M aq. citric acid (2X) and water (1X), dried over  $\text{Na}_2\text{SO}_4$ , and evaporated *in vacuo*. Compound IV is isolated by flash chromatography ( $\text{CH}_2\text{Cl}_2 \rightarrow \text{EtOAc} + \text{HOAc}$ ).

10 Step D. TentaGel S AC resin (compound V; 1 eq.) is suspended with DMF in the collection vessel of the ACT357 peptide synthesizer (Advanced ChemTech Inc. (ACT), Louisville, KY). Compound IV (3 eq.), HATU (3 eq.) and DIEA (7.5 eq.) in DMF are added and the collection vessel shaken for 1 hr. The solvent is removed and the resin washed with NMP (2X), MeOH (2X), and DMF (2X). The coupling of IV to the resin  
15 and the wash steps are repeated, to give compound VI.

Step E. The resin (compound VI) is mixed with 25% piperidine in DMF and shaken for 5 min. The resin is filtered, then mixed with 25% piperidine in DMF and shaken for 10 min. The solvent is removed and the resin washed with NMP (2X), MeOH (2X), and  
20 DMF (2X). The deprotected resin is then divided equally by the ACT357 from the collection vessel into 16 reaction vessels.

Step F. The 16 aliquots of deprotected resin from step E are suspended in DMF. To each reaction vessel is added the appropriate carboxylic acid  $\text{VII}_{1-16}$  ( $\text{R}_{1-16}\text{CO}_2\text{H}$ ; 3 eq.),  
25 HATU (3 eq.), and DIEA (7.5 eq.) in DMF. The vessels are shaken for 1 hr. The solvent is removed and the aliquots of resin washed with NMP (2X), MeOH (2X), and DMF (2X). The coupling of  $\text{VII}_{1-16}$  to the aliquots of resin and the wash steps are repeated, to give compounds  $\text{VIII}_{1-16}$ .

30 Step G. The aliquots of resin (compounds  $\text{VIII}_{1-16}$ ) are washed with  $\text{CH}_2\text{Cl}_2$  (3X). To each of the reaction vessels is added 1% TFA in  $\text{CH}_2\text{Cl}_2$  and the vessels shaken for 30 min. The solvent is filtered from the reaction vessels into individual tubes. The aliquots of resin are washed with  $\text{CH}_2\text{Cl}_2$  (2X) and MeOH (2X) and the filtrates combined into the individual tubes. The individual tubes are evaporated *in vacuo*,  
35 providing compounds  $\text{IX}_{1-16}$ .



Step II. Each of the free carboxylic acids IX<sub>1-16</sub> is dissolved in DMF. To each solution is added pyridine (1.05 eq.), followed by pentafluorophenyl trifluoroacetate (1.1 eq.). The mixtures are stirred for 45 min. at room temperature. The solutions are diluted with  
5 EtOAc, washed with 1 M aq. citric acid (3X) and 5% aq. NaHCO<sub>3</sub> (3X), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and evaporated *in vacuo*, providing compounds X<sub>1-16</sub>.

### EXAMPLE 2

#### 10 DEMONSTRATION OF PHOTOLYTIC CLEAVAGE OF T-L-X

A T-L-X compound as prepared in Example 11 was irradiated with near-UV light for 7 min at room temperature. A Rayonett fluorescence UV lamp (Southern  
15 New England Ultraviolet Co., Middletown, CT) with an emission peak at 350 nm is used as a source of UV light. The lamp is placed at a 15-cm distance from the Petri dishes with samples. SDS gel electrophoresis shows that >85% of the conjugate is cleaved under these conditions.

### EXAMPLE 3

#### 20 PREPARATION OF FLUORESCENT LABELED PRIMERS AND DEMONSTRATION OF CLEAVAGE OF FLUOROPHORE

##### Synthesis and Purification of Oligonucleotides

25 The oligonucleotides (ODNs) are prepared on automated DNA synthesizers using the standard phosphoramidite chemistry supplied by the vendor, or the H-phosphonate chemistry (Glenn Research Sterling, VA). Appropriately blocked dA, dG, dC, and T phosphoramidites are commercially available in these forms, and synthetic nucleosides may readily be converted to the appropriate form. The  
30 oligonucleotides are prepared using the standard phosphoramidite supplied by the vendor, or the H-phosphonate chemistry. Oligonucleotides are purified by adaptations of standard methods. Oligonucleotides with 5'-trityl groups are chromatographed on HPLC using a 12 micrometer, 300 # Rainin (Emeryville, CA) Dynamax C-8 4.2x250 mm reverse phase column using a gradient of 15% to 55% MeCN in 0.1 N  
35 Et<sub>3</sub>NH<sup>+</sup>OAc<sup>-</sup>, pH 7.0, over 20 min. When detritylation is performed, the

oligonucleotides are further purified by gel exclusion chromatography. Analytical checks for the quality of the oligonucleotides are conducted with a PRP-column (Alltech, Deerfield, IL) at alkaline pH and by PAGE.

Preparation of 2,4,6-trichlorotriazine derived oligonucleotides: 10 to  
5 1000  $\mu\text{g}$  of 5'-terminal amine linked oligonucleotide are reacted with an excess recrystallized cyanuric chloride in 10% n-methyl-pyrrolidone in alkaline (pH 8.3 to 8.5 preferably) buffer at 19°C to 25°C for 30 to 120 minutes. The final reaction conditions consist of 0.15 M sodium borate at pH 8.3, 2 mg/ml recrystallized cyanuric chloride and 500  $\mu\text{g}/\text{ml}$  respective oligonucleotide. The unreacted cyanuric chloride is removed by  
10 size exclusion chromatography on a G-50 Sephadex (Pharmacia, Piscataway, NJ) column.

The activated purified oligonucleotide is then reacted with a 100-fold molar excess of cystamine in 0.15 M sodium borate at pH 8.3 for 1 hour at room temperature. The unreacted cystamine is removed by size exclusion chromatography on  
15 a G-50 Sephadex column. The derived ODNs are then reacted with amine-reactive fluorochromes. The derived ODN preparation is divided into 3 portions and each portion is reacted with either (a) 20-fold molar excess of Texas Red sulfonyl chloride (Molecular Probes, Eugene, OR), with (b) 20-fold molar excess of Lissamine sulfonyl chloride (Molecular Probes, Eugene, OR), (c) 20-fold molar excess of fluorescein  
20 isothiocyanate. The final reaction conditions consist of 0.15 M sodium borate at pH 8.3 for 1 hour at room temperature. The unreacted fluorochromes are removed by size exclusion chromatography on a G-50 Sephadex column.

To cleave the fluorochrome from the oligonucleotide, the ODNs are adjusted to  $1 \times 10^{-5}$  molar and then dilutions are made (12, 3-fold dilutions) in TE (TE is  
25 0.01 M Tris, pH 7.0, 5 mM EDTA). To 100  $\mu\text{l}$  volumes of ODNs 25  $\mu\text{l}$  of 0.01 M dithiothreitol (DTT) is added. To an identical set of controls no DDT is added. The mixture is incubated for 15 minutes at room temperature. Fluorescence is measured in a black microtiter plate. The solution is removed from the incubation tubes (150 microliters) and placed in a black microtiter plate (Dynatek Laboratories, Chantilly,  
30 VA). The plates are then read directly using a Fluoroskan II fluorometer (Flow Laboratories, McLean, VA) using an excitation wavelength of 495 nm and monitoring emission at 520 nm for fluorescein, using an excitation wavelength of 591 nm and monitoring emission at 612 nm for Texas Red, and using an excitation wavelength of 570 nm and monitoring emission at 590 nm for lissamine, with the results set forth in  
35 TABLE 1.

TABLE 1

Moles of Fluorochrome	RFU non-cleaved	RFU cleaved	RFU free
$1.0 \times 10^5 \text{M}$	6.4	1200	1345
$3.3 \times 10^6 \text{M}$	2.4	451	456
$1.1 \times 10^6 \text{M}$	0.9	135	130
$3.7 \times 10^7 \text{M}$	0.3	44	48
$1.2 \times 10^7 \text{M}$	0.12	15.3	16.0
$4.1 \times 10^7 \text{M}$	0.14	4.9	5.1
$1.4 \times 10^8 \text{M}$	0.13	2.5	2.8
$4.5 \times 10^9 \text{M}$	0.12	0.8	0.9

The data indicate that there is about a 200-fold increase in relative fluorescence when  
 5 the fluorochrome is cleaved from the ODN.

#### EXAMPLE 4

##### PREPARATION OF TAGGED M13 SEQUENCE PRIMERS AND DEMONSTRATION OF CLEAVAGE OF TAGS

Preparation of 2,4,6-trichlorotriazine derived oligonucleotides: 1000  $\mu\text{g}$   
 of 5'-terminal amine linked oligonucleotide (5'-hexylamine-  
 TGTAACGACGGCCAGT-3') (Seq. ID No. 1) are reacted with an excess  
 15 recrystallized cyanuric chloride in 10% n-methyl-pyrrolidone alkaline (pH 8.3 to 8.5  
 preferably) buffer at 19 to 25- C for 30 to 120 minutes. The final reaction conditions  
 consist of 0.15 M sodium borate at pH 8.3, 2 mg/ml recrystallized cyanuric chloride and  
 500  $\mu\text{g/ml}$  respective oligonucleotide. The unreacted cyanuric chloride is removed by  
 size exclusion chromatography on a G-50 Sephadex column.

20 The activated purified oligonucleotide is then reacted with a 100-fold  
 molar excess of cystamine in 0.15 M sodium borate at pH 8.3 for 1 hour at room  
 temperature. The unreacted cystamine is removed by size exclusion chromatography on  
 a G-50 Sephadex column. The derived ODNs are then reacted with a variety of amides.  
 The derived ODN preparation is divided into 12 portions and each portion is reacted (25  
 25 molar excess) with the pentafluorophenyl-esters of either: (1) 4-methoxybenzoic acid,

(2) 4-fluorobenzoic acid, (3) toluic acid, (4) benzoic acid, (5) indole-3-acetic acid, (6) 2,6-difluorobenzoic acid, (7) nicotinic acid N-oxide, (8) 2-nitrobenzoic acid, (9) 5-acetylsalicylic acid, (10) 4-ethoxybenzoic acid, (11) cinnamic acid, (12) 3-aminonicotinic acid. The reaction is for 2 hours at 37°C in 0.2 M NaBorate pH 8.3.

5 The derived ODNs are purified by gel exclusion chromatography on G-50 Sephadex.

To cleave the tag from the oligonucleotide, the ODNs are adjusted to  $1 \times 10^{-5}$  molar and then dilutions are made (12, 3-fold dilutions) in TE (TE is 0.01 M Tris, pH 7.0, 5 mM EDTA) with 50% EtOH (V/V). To 100  $\mu$ l volumes of ODNs 25  $\mu$ l of 0.01 M dithiothreitol (DTT) is added. To an identical set of controls no DDT is added.

10 Incubation is for 30 minutes at room temperature. NaCl is then added to 0.1 M and 2 volumes of EtOH is added to precipitate the ODNs. The ODNs are removed from solution by centrifugation at 14,000 x G at 4°C for 15 minutes. The supernatants are reserved, dried to completeness. The pellet is then dissolved in 25  $\mu$ l MeOH. The pellet is then tested by mass spectrometry for the presence of tags.

15 The mass spectrometer used in this work is an external ion source Fourier-transform mass spectrometer (FTMS). Samples prepared for MALDI analysis are deposited on the tip of a direct probe and inserted into the ion source. When the sample is irradiated with a laser pulse, ions are extracted from the source and passed into a long quadrupole ion guide that focuses and transports them to an FTMS analyzer  
20 cell located inside the bore of a superconducting magnet.

The spectra yield the following information. Peaks varying in intensity from 25 to 100 relative intensity units at the following molecular weights: (1) 212.1 amu indicating 4-methoxybenzoic acid derivative, (2) 200.1 indicating 4-fluorobenzoic acid derivative, (3) 196.1 amu indicating toluic acid derivative, (4) 182.1 amu indicating  
25 benzoic acid derivative, (5) 235.2 amu indicating indole-3-acetic acid derivative, (6) 218.1 amu indicating 2,6-difluorobenzoic derivative, (7) 199.1 amu indicating nicotinic acid N-oxide derivative, (8) 227.1 amu indicating 2-nitrobenzamide, (9) 179.18 amu indicating 5-acetylsalicylic acid derivative, (10) 226.1 amu indicating 4-ethoxybenzoic acid derivative, (11) 209.1 amu indicating cinnamic acid derivative,  
30 (12) 198.1 amu indicating 3-aminonicotinic acid derivative.

The results indicate that the MW-identifiers are cleaved from the primers and are detectable by mass spectrometry.

EXAMPLE 5PREPARATION OF A SET OF COMPOUNDS  
OF THE FORMULA  $R_{1-36}$ -LYS( $\epsilon$ -INIP)-ANP-TFP

5           Figure 3 illustrates the parallel synthesis of a set of 36 T-L-X compounds ( $X = L_h$ ), where  $L_h$  is an activated ester (specifically, tetrafluorophenyl ester),  $L^2$  is an ortho-nitrobenzylamine group with  $L^3$  being a methylene group that links  $L_h$  and  $L^2$ , T has a modular structure wherein the carboxylic acid group of lysine has been joined to the nitrogen atom of the  $L^2$  benzylamine group to form an amide bond, and a variable  
10 weight component  $R_{1-36}$ , (where these R groups correspond to  $T^2$  as defined herein, and may be introduced via any of the specific carboxylic acids listed herein) is bonded through the  $\alpha$ -amino group of the lysine, while a mass spec sensitivity enhancer group (introduced via N-methylisonipecotic acid) is bonded through the  $\epsilon$ -amino group of the lysine.

15           Referring to Figure 3:

Step A. NovaSyn HMP Resin (available from NovaBiochem; 1 eq.) is suspended with DMF in the collection vessel of the ACT357. Compound I (ANP available from ACT; 3 eq.), HATU (3 eq.) and NMM (7.5 eq.) in DMF are added and the collection vessel shaken for 1 hr. The solvent is removed and the resin washed with NMP (2X), MeOH  
20 (2X), and DMF (2X). The coupling of I to the resin and the wash steps are repeated, to give compound II.

Step B. The resin (compound II) is mixed with 25% piperidine in DMF and shaken for 5 min. The resin is filtered, then mixed with 25% piperidine in DMF and shaken for 10  
25 min. The solvent is removed, the resin washed with NMP (2X), MeOH (2X), and DMF (2X), and used directly in step C.

Step C. The deprotected resin from step B is suspended in DMF and to it is added an Fmoc-protected amino acid, containing a protected amine functionality in its side  
30 chain (Fmoc-Lysine(Aloc)-OH, available from PerSeptive Biosystems; 3 eq.), HATU (3 eq.), and NMM (7.5 eq.) in DMF. The vessel is shaken for 1 hr. The solvent is removed and the resin washed with NMP (2X), MeOH (2X), and DMF (2X). The coupling of Fmoc-Lys(Aloc)-OH to the resin and the wash steps are repeated, to give compound IV.

35

Step D. The resin (compound IV) is washed with  $\text{CH}_2\text{Cl}_2$  (2X), and then suspended in a solution of  $(\text{PPh}_3)_4\text{Pd}$  (0) (0.3 eq.) and  $\text{PhSiH}_3$  (10 eq.) in  $\text{CH}_2\text{Cl}_2$ . The mixture is shaken for 1 hr. The solvent is removed and the resin is washed with  $\text{CH}_2\text{Cl}_2$  (2X). The palladium step is repeated. The solvent is removed and the resin is washed with  $\text{CH}_2\text{Cl}_2$  (2X), N,N-diisopropylethylammonium diethyldithiocarbamate in DMF (2X), DMF (2X) to give compound V.

Step E. The deprotected resin from step D is coupled with N-methylisonipecotic acid as described in step C to give compound VI.

Step F. The Fmoc protected resin VI is divided equally by the ACT357 from the collection vessel into 36 reaction vessels to give compounds  $\text{VI}_{1-36}$ .

Step G. The resin (compounds  $\text{VI}_{1-36}$ ) is treated with piperidine as described in step B to remove the Fmoc group.

Step H. The 36 aliquots of deprotected resin from step G are suspended in DMF. To each reaction vessel is added the appropriate carboxylic acid ( $\text{R}_{1-36}\text{CO}_2\text{H}$ ; 3 eq.), HATU (3 eq.), and NMM (7.5 eq.) in DMF. The vessels are shaken for 1 hr. The solvent is removed and the aliquots of resin washed with NMP (2X), MeOH (2X), and DMF (2X). The coupling of  $\text{R}_{1-36}\text{CO}_2\text{H}$  to the aliquots of resin and the wash steps are repeated, to give compounds  $\text{VIII}_{1-36}$ .

Step I. The aliquots of resin (compounds  $\text{VIII}_{1-36}$ ) are washed with  $\text{CH}_2\text{Cl}_2$  (3X). To each of the reaction vessels is added 90:5:5 TFA:H<sub>2</sub>O: $\text{CH}_2\text{Cl}_2$  and the vessels shaken for 120 min. The solvent is filtered from the reaction vessels into individual tubes. The aliquots of resin are washed with  $\text{CH}_2\text{Cl}_2$  (2X) and MeOH (2X) and the filtrates combined into the individual tubes. The individual tubes are evaporated *in vacuo*, providing compounds  $\text{IX}_{1-36}$ .

Step J. Each of the free carboxylic acids  $\text{IX}_{1-36}$  is dissolved in DMF. To each solution is added pyridine (1.05 eq.), followed by tetrafluorophenyl trifluoroacetate (1.1 eq.). The mixtures are stirred for 45 min. at room temperature. The solutions are diluted with EtOAc, washed with 5% aq.  $\text{NaHCO}_3$  (3X), dried over  $\text{Na}_2\text{SO}_4$ , filtered, and evaporated *in vacuo*, providing compounds  $\text{X}_{1-36}$ .

### EXAMPLE 6

#### PREPARATION OF A SET OF COMPOUNDS OF THE FORMULA $R_{1-36}$ -LYS( $\epsilon$ -INIP)-NBA-TFP

5

Figure 4 illustrates the parallel synthesis of a set of 36 T-L-X compounds ( $X = L_h$ ), where  $L_h$  is an activated ester (specifically, tetrafluorophenyl ester),  $L^2$  is an ortho-nitrobenzylamine group with  $L^3$  being a direct bond between  $L_h$  and  $L^2$ , where  $L_h$  is joined directly to the aromatic ring of the  $L^2$  group, T has a modular structure wherein the carboxylic acid group of lysine has been joined to the nitrogen atom of the  $L^2$  benzylamine group to form an amide bond, and a variable weight component  $R_{1-36}$ , (where these R groups correspond to  $T^2$  as defined herein, and may be introduced via any of the specific carboxylic acids listed herein) is bonded through the  $\alpha$ -amino group of the lysine, while a mass spec enhancer group (introduced via N-methylisonipecotic acid) is bonded through the  $\epsilon$ -amino group of the lysine.

15

Referring to Figure 4

Step A. NovaSyn HMP Resin is coupled with compound I (NBA prepared according to the procedure of Brown et al., Molecular Diversity, 1, 4 (1995)) according to the procedure described in step A of Example 5, to give compound II.

20

Steps B-J. The resin (compound II) is treated as described in steps B-J of Example 5 to give compounds  $X_{1-36}$ .

### EXAMPLE 7

#### PREPARATION OF A SET OF COMPOUNDS OF THE FORMULA INIP-LYS ( $\epsilon$ - $R_{1-36}$ )-ANP-TFP

25

Figure 5 illustrates the parallel synthesis of a set of 36 T-L-X compounds ( $X = L_h$ ), where  $L_h$  is an activated ester (specifically, tetrafluorophenyl ester),  $L^2$  is an ortho-nitrobenzylamine group with  $L^3$  being a methylene group that links  $L_h$  and  $L^2$ , T has a modular structure wherein the carboxylic acid group of lysine has been joined to the nitrogen atom of the  $L^2$  benzylamine group to form an amide bond, and a variable weight component  $R_{1-36}$ , (where these R groups correspond to  $T^2$  as defined herein, and may be introduced via any of the specific carboxylic acids listed herein) is bonded through the  $\epsilon$ -amino group of the lysine, while a mass spec sensitivity enhancer group

30

35

(introduced via N-methylisonipecotic acid) is bonded through the  $\alpha$ -amino group of the lysine.

Referring to Figure 5:

Steps A-C. Same as in Example 5.

5

Step D. The resin (compound IV) is treated with piperidine as described in step B of Example 5 to remove the Fmoc group.

10 Step E. The deprotected  $\alpha$ -amine on the resin in step D is coupled with N-methylisonipecotic acid as described in step C of Example 5 to give compound V.

Step F. Same as in Example 5.

15 Step G. The resin (compounds VI<sub>1-36</sub>) are treated with palladium as described in step D of Example 5 to remove the Alloc group.

Steps H-J. The compounds X<sub>1-36</sub> are prepared in the same manner as in Example 5.

#### EXAMPLE 8

20

#### PREPARATION OF A SET OF COMPOUNDS OF THE FORMULA R<sub>1-36</sub>-GLU( $\gamma$ -DIAEA)-ANP-TFP

Figure 6 illustrates the parallel synthesis of a set of 36 T-L-X compounds (X = L<sub>h</sub>), where L<sub>h</sub> is an activated ester (specifically, tetrafluorophenyl ester), L<sup>2</sup> is an ortho-nitrobenzylamine group with L<sup>3</sup> being a methylene group that links L<sub>h</sub> and L<sup>2</sup>, T has a modular structure wherein the  $\alpha$ -carboxylic acid group of glutamic acid has been joined to the nitrogen atom of the L<sup>2</sup> benzylamine group to form an amide bond, and a variable weight component R<sub>1-36</sub>, (where these R groups correspond to T<sup>2</sup> as defined herein, and may be introduced via any of the specific carboxylic acids listed  
25 herein) is bonded through the  $\alpha$ -amino group of the glutamic acid, while a mass spec sensitivity enhancer group (introduced via 2-(diisopropylamino)ethylamine) is bonded  
30 through the  $\gamma$ -carboxylic acid of the glutamic acid.

Referring to Figure 6:

Steps A-B. Same as in Example 5.

35



Step C. The deprotected resin (compound III) is coupled to Fmoc-Glu-(OAl)-OH using the coupling method described in step C of Example 5 to give compound IV.

Step D. The allyl ester on the resin (compound IV) is washed with  $\text{CH}_2\text{Cl}_2$  (2X) and mixed with a solution of  $(\text{PPh}_3)_4\text{Pd}$  (0) (0.3 eq.) and N-methylaniline (3 eq.) in  $\text{CH}_2\text{Cl}_2$ . The mixture is shaken for 1 hr. The solvent is removed and the resin is washed with  $\text{CH}_2\text{Cl}_2$  (2X). The palladium step is repeated. The solvent is removed and the resin is washed with  $\text{CH}_2\text{Cl}_2$  (2X), N,N-diisopropylethylammonium diethyldithiocarbamate in DMF (2X), DMF (2X) to give compound V.

Step E. The deprotected resin from step D is suspended in DMF and activated by mixing HATU (3 eq.), and NMM (7.5 eq.). The vessels are shaken for 15 minutes. The solvent is removed and the resin washed with NMP (1X). The resin is mixed with 2-(diisopropylamino)ethylamine (3 eq.) and NMM (7.5 eq.). The vessels are shaken for 1 hour. The coupling of 2-(diisopropylamino)ethylamine to the resin and the wash steps are repeated, to give compound VI.

Steps F-J. Same as in Example 5.

## EXAMPLE 9

### PREPARATION OF A SET OF COMPOUNDS OF THE FORMULA $\text{R}_{1-36}\text{-LYS}(\epsilon\text{-INIP})\text{-ANP-LYS}(\epsilon\text{-NH}_2)\text{-NH}_2$

Figure 7 illustrates the parallel synthesis of a set of 36 T-L-X compounds ( $\text{X} = \text{L}_h$ ), where  $\text{L}_h$  is an amine (specifically, the  $\epsilon$ -amino group of a lysine-derived moiety),  $\text{L}^2$  is an ortho-nitrobenzylamine group with  $\text{L}^3$  being a carboxamido-substituted alkyleneaminoacylalkylene group that links  $\text{L}_h$  and  $\text{L}^2$ , T has a modular structure wherein the carboxylic acid group of lysine has been joined to the nitrogen atom of the  $\text{L}^2$  benzylamine group to form an amide bond, and a variable weight component  $\text{R}_{1-36}$ , (where these R groups correspond to  $\text{T}^2$  as defined herein, and may be introduced via any of the specific carboxylic acids listed herein) is bonded through the  $\alpha$ -amino group of the lysine, while a mass spec sensitivity enhancer group (introduced via N-methylisonipecotic acid) is bonded through the  $\epsilon$ -amino group of the lysine.

Referring to Figure 7:

Step A. Fmoc-Lys(Boc)-SRAM Resin (available from ACT; compound I) is mixed with 25% piperidine in DMF and shaken for 5 min. The resin is filtered, then mixed with 25% piperidine in DMF and shaken for 10 min. The solvent is removed, the resin washed with NMP (2X), MeOH (2X), and DMF (2X), and used directly in step B.

Step B. The resin (compound II), ANP (available from ACT; 3 eq.), HATU (3 eq.) and NMM (7.5 eq.) in DMF are added and the collection vessel shaken for 1 hr. The solvent is removed and the resin washed with NMP (2X), MeOH (2X), and DMF (2X). The coupling of I to the resin and the wash steps are repeated, to give compound III.

Steps C-J. The resin (compound III) is treated as in steps B-I in Example 5 to give compounds  $X_{1-36}$ .

#### EXAMPLE 10

##### PREPARATION OF A SET OF COMPOUNDS OF THE FORMULA $R_{1-36}$ -LYS( $\epsilon$ -TFA)-LYS( $\epsilon$ -IINP)-ANP-TFP

Figure 8 illustrates the parallel synthesis of a set of 36 T-L-X compounds ( $X = L_h$ ), where  $L_h$  is an activated ester (specifically, tetrafluorophenyl ester),  $L^2$  is an ortho-nitrobenzylamine group with  $L^3$  being a methylene group that links  $L_h$  and  $L^2$ , T has a modular structure wherein the carboxylic acid group of a first lysine has been joined to the nitrogen atom of the  $L^2$  benzylamine group to form an amide bond, a mass spec sensitivity enhancer group (introduced via N-methylisonipecotic acid) is bonded through the  $\epsilon$ -amino group of the first lysine, a second lysine molecule has been joined to the first lysine through the  $\alpha$ -amino group of the first lysine, a molecular weight adjuster group (having a trifluoroacetyl structure) is bonded through the  $\epsilon$ -amino group of the second lysine, and a variable weight component  $R_{1-36}$ , (where these R groups correspond to  $T^2$  as defined herein, and may be introduced via any of the specific carboxylic acids listed herein) is bonded through the  $\alpha$ -amino group of the second lysine. Referring to Figure 8:

Steps A-E. These steps are identical to steps A-E in Example 5.

Step F. The resin (compound VI) is treated with piperidine as described in step B in Example 5 to remove the Fmoc group.

Step G. The deprotected resin (compound VII) is coupled to Fmoc-Lys(Tfa)-OH using the coupling method described in step C of Example 5 to give compound VIII.

Steps H-K. The resin (compound VIII) is treated as in steps F-J in Example 5 to give compounds XI<sub>1-36</sub>.

### EXAMPLE 11

#### PREPARATION OF A SET OF COMPOUNDS OF THE FORMULA R<sub>1-36</sub>-LYS(ε-INIP)-ANP-5'-AH-ODN

Figure 9 illustrates the parallel synthesis of a set of 36 T-L-X compounds (X = MOI, where MOI is a nucleic acid fragment, ODN) derived from the esters of Example 5 (the same procedure could be used with other T-L-X compounds wherein X is an activated ester). The MOI is conjugated to T-L through the 5' end of the MOI, via a phosphodiester - alkyleneamine group.

Referring to Figure 9:

Step A. Compounds XII<sub>1-36</sub> are prepared according to a modified biotinylation procedure in Van Ness et al., Nucleic Acids Res., 19, 3345 (1991). To a solution of one of the 5'-aminohexyl oligonucleotides (compounds XI<sub>1-36</sub>, 1 mg) in 200 mM sodium borate (pH 8.3, 250 mL) is added one of the Tetrafluorophenyl esters (compounds X<sub>1-36</sub> from Example A, 100-fold molar excess in 250 mL of NMP). The reaction is incubated overnight at ambient temperature. The unreacted and hydrolyzed tetrafluorophenyl esters are removed from the compounds XII<sub>1-36</sub> by Sephadex G-50 chromatography.

### EXAMPLE 12

#### PREPARATION OF A SET OF COMPOUNDS OF THE FORMULA R<sub>1-36</sub>-LYS(ε-INIP)-ANP-LYS(ε-(MCT-5'-AH-ODN))-NH<sub>2</sub>

Figure 10 illustrates the parallel synthesis of a set of 36 T-L-X compounds (X = MOI, where MOI is a nucleic acid fragment, ODN) derived from the amines of Example 11 (the same procedure could be used with other T-L-X compounds wherein X is an amine). The MOI is conjugated to T-L through the 5' end of the MOI, via a phosphodiester - alkyleneamine group.

Referring to Figure 10:

Step A. The 5'-[6-(4,6-dichloro-1,3,5-triazin-2-ylamino)hexyl]oligonucleotides XII<sub>1-36</sub> are prepared as described in Van Ness et al., Nucleic Acids Res., 19, 3345 (1991).

Step B. To a solution of one of the 5'-[6-(4,6-dichloro-1,3,5-triazin-2-ylamino)hexyl]oligonucleotides (compounds XII<sub>1-36</sub>) at a concentration of 1 mg/ml in 100 mM sodium borate (pH 8.3) was added a 100-fold molar excess of a primary amine selected from R<sub>1-36</sub>-Lys(e-iNIP)-ANP-Lys(e-NH<sub>2</sub>)-NH<sub>2</sub> (compounds X<sub>1-36</sub> from Example 11). The solution is mixed overnight at ambient temperature. The unreacted amine is removed by ultrafiltration through a 3000 MW cutoff membrane (Amicon, Beverly, MA) using H<sub>2</sub>O as the wash solution (3 X). The compounds XIII<sub>1-36</sub> are isolated by reduction of the volume to 100 mL.

### EXAMPLE 13

#### DEMONSTRATION OF THE SIMULTANEOUS DETECTION OF MULTIPLE TAGS BY MASS SPECTROMETRY

This example provides a description of the ability to simultaneously detect multiple compounds (tags) by mass spectrometry. In this particular example, 31 compounds are mixed with a matrix, deposited and dried on to a solid support and then desorbed with a laser. The resultant ions are then introduced in a mass spectrometer.

The following compounds (purchased from Aldrich, Milwaukee, WI) are mixed together on an equal molar basis to a final concentration of 0.002 M (on a per compound) basis: benzamide (121.14), nicotinamide (122.13), pyrazinamide (123.12), 3-amino-4-pyrazolecarboxylic acid (127.10), 2-thiophenecarboxamide (127.17), 4-aminobenzamide (135.15), toluamide (135.17), 6-methylnicotinamide (136.15), 3-aminonicotinamide (137.14), nicotinamide N-oxide (138.12), 3-hydropicolinamide (138.13), 4-fluorobenzamide (139.13), cinnamamide (147.18), 4-methoxybenzamide (151.17), 2,6-difluorobenzamide (157.12), 4-amino-5-imidazole-carboxamide (162.58), 3,4-pyridine-dicarboxamide (165.16), 4-ethoxybenzamide (165.19), 2,3-pyrazinedicarboxamide (166.14), 2-nitrobenzamide (166.14), 3-fluoro-4-methoxybenzoic acid (170.4), indole-3-acetamide (174.2), 5-acetylsalicylamide (179.18), 3,5-dimethoxybenzamide (181.19), 1-naphthaleneacetamide (185.23), 8-chloro-3,5-diamino-2-pyrazinecarboxamide (187.59), 4-trifluoromethyl-benzamide (189.00), 5-amino-5-phenyl-4-pyrazole-carboxamide (202.22), 1-methyl-2-benzyl-malonamate (207.33), 4-amino-2,3,5,6-tetrafluorobenzamide (208.11), 2,3-

naphthenedicarboxylic acid (212.22). The compounds are placed in DMSO at the concentration described above. One  $\mu$ l of the material is then mixed with alpha-cyano-4-hydroxy cinnamic acid matrix (after a 1:10,000 dilution) and deposited on to a solid stainless steel support.

5                   The material is then desorbed by a laser using the Protein TOF Mass Spectrometer (Bruker, Manning Park, MA) and the resulting ions are measured in both the linear and reflectron modes of operation. The following m/z values are observed (Figure 11):

10	121.1---->	benzamide (121.14)
	122.1---->	nicotinamide (122.13)
	123.1---->	pyrazinamide (123.12)
	124.1	
	125.2	
15	127.3---->	3-amino-4-pyrazolecarboxylic acid (127.10)
	127.2---->	2-thiophenecarboxamide (127.17)
	135.1---->	4-aminobenzamide (135.15)
	135.1---->	tolumide (135.17)
	136.2---->	6-methylnicotinamide (136.15)
20	137.1---->	3-aminonicotinamide (137.14)
	138,2---->	nicotinamide N-oxide (138.12)
	138.2---->	3-hydropicolinamide (138.13)
	139.2---->	4-fluorobenzamide (139.13)
	140.2	
25	147.3---->	cinnamamide (147.18)
	148.2	
	149.2	
		4-methoxybenzamide (151.17)
	152.2	
30		2,6-difluorbenzamide (157.12)
	158.3	
		4-amino-5-imidazole-carboxamide (162.58)
	163.3	
	165.2---->	3,4-pyridine-dicarboxamide (165.16)
35	165.2---->	4-ethoxybenzamide (165.19)

181

	166.2----->	2,3-pyrazinedicarboxamide (166.14)
	166.2----->	2-nitrobenzamide (166.14)
		3-fluoro-4-methoxybenzoic acid (170.4)
	171.1	
5	172.2	
	173.4	
		indole-3-acetamide (174.2)
	178.3	
	179.3----->	5-acetylsalicylamide (179.18)
10	181.2----->	3,5-dimethoxybenzamide (181.19)
	182.2----->	
		1-naphthaleneacetamide (185.23)
	186.2	
		8-chloro-3,5-diamino-2-pyrazinecarboxamide (187.59)
15	188.2	
	189.2----->	4-trifluoromethyl-benzamide (189.00)
	190.2	
	191.2	
	192.3	
20		5-amino-5-phenyl-4-pyrazole-carboxamide (202.22)
	203.2	
	203.4	
		1-methyl-2-benzyl-malonamate (207.33)
		4-amino-2,3,5,6-tetrafluorobenzamide (208.11)
25	212.2----->	2,3-naphthlenedicarboxylic acid (212.22).
	219.3	
	221.2	
	228.2	
	234.2	
30	237.4	
	241.4	

The data indicate that 22 of 31 compounds appeared in the spectrum with the anticipated mass, 9 of 31 compounds appeared in the spectrum with a n + H mass (1 atomic mass unit, amu) over the anticipated mass. The latter phenomenon is probably

due to the protonation of an amine within the compounds. Therefore 31 of 31 compounds are detected by MALDI Mass Spectroscopy. More importantly, the example demonstrates that multiple tags can be detected simultaneously by a spectroscopic method.

5           The alpha-cyano matrix alone (Figure 11) gave peaks at 146.2, 164.1, 172.1, 173.1, 189.1, 190.1, 191.1, 192.1, 212.1, 224.1, 228.0, 234.3. Other identified masses in the spectrum are due to contaminants in the purchased compounds as no effort was made to further purify the compounds.

10

#### EXAMPLE 14

##### MICROSATELLITE MARKERS: PCR AMPLIFICATIONS.

15           The microsatellite markers are amplified utilizing the following standard PCR conditions. Briefly, PCR reactions are performed in a total volume of 50  $\mu$ l, containing 40 ng of genomic DNA, 50 pmol of each primer, 0.125 mM dNTPs and 1 unit of *Taq* polymerase. 1X amplification buffer contains 10 mM Tris base, pH 9, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.1% Triton X-100 and 0.01% gelatin. The reactions are performed using a "hot-start" procedure: *Taq* polymerase is added only after a first  
20   denaturation step of 5 minutes at 96°C. Amplification is carried out for 35 cycles: denaturation (94°C for 40 sec) and annealing (55°C for 30 sec). An elongation step (72°C for 2 minutes) ends the process after the last annealing. Since the amplification products to be obtained are short (90 to 350 base pairs long) and the time interval to raise the temperature from 55°C to 94°C (obtained with a ramping rate of 1°C/second)  
25   is long enough, completion of DNA elongation can be achieved without a step at 72°C.

#### EXAMPLE 15

##### SEPARATION OF DNA FRAGMENTS

30

##### Instrumentation

35           The separation of DNA fragments can be performed using an HPLC system assembled from several standard components. These components include a minimum of two high pressure pumps which pump solvent through a high pressure gradient mixer, an injector, HPLC column, and a detector. The injector is an

automated, programmable autosampler capable of storing typically between eighty and one hundred samples at or below ambient temperatures to maintain the stability of the sample components. The autoinjector also is capable of making uL size injections in a reproducible manner completely unattended. The HPLC column is contained in a  
5 heated column compartment capable of holding a defined temperature to within 0.1°C. The column used in the examples below was purchased from SeraSep (San Jose, CA) under the name DNASep. This column is a 55x4.6 mm column with a 2.2um non-porous polystyrenedivinylbenzene copolymer particle alkylated with C18. The packing material is stable within a pH range of 2-12 and tolerates temperatures as high as 70°C.  
10 Detection of analyte was performed using a single or multiple wavelength UV detector or diode array detector.

### Methods

The methods applied in this example for separation of DNA fragments  
15 use ion-pair chromatography, a form of chromatography in which ions in solution can be paired or neutralized and separated as an ion pair on a reversed phase column. The lipophilic character and the concentration of the counterion determine the degree of retention of the analyte. In the case of a DNA molecule the lipophilic, cationic buffer component pairs with anionic phosphate groups of the DNA backbone. The buffer  
20 components also interact with the alkyl groups of the stationary phase. The paired DNA then elutes according to size as the mobile phase is made progressively more organic with increasing concentration of acetonitrile. Evaluation of the suitability of various amine salts was evaluated using enzymatic digests of plasmids or commercially available DNA ladders. The range of acetonitrile required to elute the DNA as well as  
25 the temperature of the column compartment varied with each buffer evaluated.

### Buffers

The buffers evaluated for their ion-pairing capability were prepared from stock solutions. In order to keep the concentration of ion-pair reagent the same  
30 throughout the gradient, the ion-pair reagent was added to both the water and the acetonitrile mobile phases. The column was equilibrated with a new mobile phase for approximately 18 hours at a flow rate of 50ul/minute before attempting any separation. Once a mobile phase had been evaluated, it was removed and the column cleaned with a flush of 800 mL 0.1% formic acid in 50% acetonitrile, followed by a flush with 800 mL  
35 0.1% acetic acid in 50% acetonitrile before equilibration with a new mobile phase.



A. nn-Dimethyloctylammonium trifluoroacetate

A stock solution of 1molar dimethyloctylammonium trifluoroacetate was prepared by mixing one half of an equivalent of trifluoroacetic acid in an appropriate volume of water and slowly adding one equivalent of nn-Dimethyloctylamine. The pH of this stock solution is 7. The stock solution was diluted with an appropriate volume of water or acetonitrile to working concentration.

B. nn-Dimethylheptylammonium acetate

A stock solution of 1molar dimethylheptylammonium acetate was prepared by mixing one equivalent of glacial acetic acid in an appropriate volume of water and slowly adding one equivalent of nn-Dimethylheptylamine. The pH of this stock solution is 6.6. The stock solution was diluted with an appropriate amount of water or acetonitrile to working concentration.

C. nn-Dimethylhexylammonium acetate

A stock solution of 1 molar dimethylhexylammonium acetate was prepared by mixing one equivalent of glacial acetic acid in an appropriate volume of water and slowly adding one equivalent of nn-Dimethylhexylamine. The pH of this stock solution is 6.5. The stock solution was diluted with an appropriate volume of water or acetonitrile to working concentration.

D. nn-Dimethylbutylammonium acetate

A stock solution of 1molar dimethylbutylammonium acetate was prepared by mixing one equivalent of glacial acetic acid in an appropriate volume of water and slowly adding one equivalent of nn-Dimethylbutylamine. The pH of the stock solution is 6.9. The stock solution was diluted with an appropriate volume of water or acetonitrile to working concentration.

E. nn-Dimethylisopropylammonium acetate

A stock solution of 1 molar dimethylisopropylammonium acetate was prepared by mixing one equivalent of glacial acetic acid in an appropriate volume of water and slowly adding one equivalent of nn-Dimethylisopropylamine. The pH of the stock solution is 6.9. The stock solution was diluted with an appropriate volume of water or acetonitrile to working concentration.

F. nn-Dimethylcyclohexylammonium acetate

A stock solution of 1 molar dimethylcyclohexylammonium acetate was prepared by mixing one equivalent of glacial acetic acid in appropriate volume of water and slowly adding one equivalent of nn-Dimethylcyclohexylamine. The pH of the stock solution is 6.5. The stock solution was diluted with an appropriate volume of water or acetonitrile to working concentration.

G. Methylpiperidine acetate

A stock solution of 1 molar methylpiperidine acetate was prepared by mixing one equivalent of glacial acetic acid in an appropriate volume of water and slowly adding one equivalent of 1-methylpiperidine. The pH of the solution is 7. The stock solution was diluted with an appropriate volume of water or acetonitrile to working concentration.

H. Methylpyrrolidine acetate

A stock solution of 1 molar piperidine acetate was prepared by mixing one equivalent of glacial acetic acid in an appropriate volume of water and slowly adding one equivalent 1-methylpyrrolidine. The pH of the stock solution is 7. The stock solution was diluted in an appropriate volume of water or acetonitrile to working concentration.

I. Triethylammonium acetate

A stock solution of 2 molar triethylammonium acetate pH 7.0 was purchased from Glenn Research Sterling, Virginia. The stock solution was diluted in an appropriate volume of water or acetonitrile to working concentration.

EXAMPLE 16

## DNA FINGERPRINT

DNA fingerprinting adaptors are prepared comprising the following: a core sequence and an enzyme specific sequence. The structure of the *Eco*R1-adaptor is 5'-CTCGTAGACTGCGTACC (SEQ ID No. \_\_\_\_), the structure of the *Mse*I-adaptor is: 5"-GACGATGAGTCCTGAG

Adapters for the rare cutter enzymes were identical to the *EcoRI* with the exception that cohesive ends were used. ALPH primers consists of three parts: a core sequence, an enzyme specific sequence and a selective extension sequence. The *EcoRI* and *MseI* primers are described as follows: *EcoRI*: 5'-gactgcgtaaa-aattc-NNN (SEQ ID No. \_\_\_\_); *MseI*: 5'-gatgagtcctgag-taa-NNN (SEQ ID No. \_\_\_\_).

Genomic DNA was incubated for 1 hour at 37°C with 5 units *EcoRI* and 5 units of *MseI* in 40 µl volumes with 10 mM Tris-acetate pH 7.5, 10 mM MgAce, 50 mM KAcetate, 5 mM DTT, 50 ng/microliter BSA, 5 mM DTT. Next, 10 µl of a solution containing 5 pMol *EcoRI* adapters, 50 pMol *MseI* adapters,, 1 unit of T4 ligase, 1 mM ATP, in 10 mM Tris-acetate pH 7.5, 10 mM MgAce, 50 mM KAcetate, 5 mM DTT, 50 ng/microliter BSA was added and the incubation was continued for 3 hours at 37°C. Adapters were prepared by adding equimolar amounts of both strands: adapters were not phosphorylated. After ligation, the reaction mixture was diluted to 500 µl with 10 mM Tris HCl, 0.1 mM EDTA pH8.0 and stored at -20°C.

Genetic fingerprinting reactions: Amplification reactions are described using DNA templates for the enzyme combination *EcoRI/MseI*. Genomic fingerprints with other enzyme combinations were performed with appropriate primers. The amplification reactions generally employed two oligonucleotides, one corresponding to the *EcoRI* pends and one corresponding to the *MseI* -ends. One of the two primers was labelled with the CMST tag, preferably the *ECORI* primer. The PCR s were performed using 5 ng labeled *EcoRI* primer, 30 ng *MseI* primer, 5 microliters of template DNA, 0.4 units Taq polymerase, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCL2, 50 mM KCl, 0.2 mM of dATP, dGTP, dCTP, dTTP. The PCR reactions differed depending on the nature of the selective amplification extensions of the DNA fingerprinting primers used for amplification. DNA fingerprinting reactions with primers having two two or three selective nucleotides were performed for 36 cycles with the following cycle profile: a 30 second DNA denaturation step, at 94°C, a 30 second annealing step at 55°C, and then a 1 minute extension step at 72°C for 1 minute. The annealing temperature in the first step was 65°C and was subsequently reduce for each cycle step by 0.7°C in the next 12 cycles and was continued at 56°C for the remaining 23 cycles. All amplifications were performed in a an MJ thermocycler (Watertown MA).

DNA fingerprinting of the complex genomes (such as humans) involve two amplification steps. The preamplification was performed with two DNA fingerprinting having a single selective nucleotide as described above with the exception that 30 ng of both DNA fingerprinting primers was used and that these

primers were not labelled with CMST, after the preamplification step, the reaction mixtures were diluted 10-fold with with 10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0, and used as templates for the second amplification reaction. The second amplification reaction was performed as described above for DNA fingerprinting reactions with  
5 primers having the longer selective extensions.

The products from the amplification reactions were analyzed by HPLC. HPLC was carried out using automated HPLC instrumentation (Rainin, Emeryville, CA., or, Hewlett Packard, Palo Alto, CA). Unpurified DNA fingerprinting products which had been denatured for 3 minutes at 95 prior into injection into an HPLC were  
10 eluted with linear acetonitrile (ACN, J.T. Baker, NJ) gradient of 1.8%/minute at a flow rate of 0.9 ml/minute, The start and end points were adjusted according to the size of the amplified products. The temperature required for the successful resolution of the molecules generated during the DNA fingerprinting technique was 50°C. The effluent from the HPLC was then directed into a mass spectrometer (Hewlett Packard, Palo  
15 Alto, CA) for the detection of tags.

The following fragments eluted in the order presented (The number sited are the positions within the lambda genome at which a cleavage site occurred): 47, 78, 91, 733, 1456, 2176, 3275, 3419, 4349, 444, 5268, 5709, 6076, 6184, 6551, 7024, 7949, 8062, 8200, 8461, 9079, 9253, 9692, 9952, 11083, 11116, 11518, 11584, 12619, 12967,  
20 14108, 14892, 15628, 15968, 16034, 16295, 16859, 18869, 19137, 19482, 20800, 21226, 21441, 2635, 21702, 21903, 21948, 22724, 23048, 23084, 23111, 23206, 23279, 23285, 23479, 23498, 23555, 23693, 23887, 23979, 23987, 24073, 24102, 24751, 24987, 25170, 25255, 25353, 25437, 26104, 25578, 25746, 25968, 26133, 26426, 26451, 26483, 26523, 26585, 26651, 26666, 26679, 26693, 26763, 26810, 26984,  
25 26993, 27038, 27092, 27203, 27317, 27683, 28456, 28569, 28922, 28972, 29374, 29981, 30822, 30620, 30639, 30722, 30735, 30756, 31169, 31747, 31808, 32194, 32218, 32641, 32704, 33222, 33351, 33688, 33736, 33748, 33801, 34202, 34366, 34406, 34590, 34618, 34684, 34735, 34753, 34831, 35062, 35269, 35534, 35541, 36275, 36282, 36303, 36430, 36492, 36531, 36543, 36604, 36736, 36757, 36879,  
30 37032, 37442, 37766, 37783, 37882, 37916, 37994, 36164, 38287, 38412, 38834, 39168, 44972, 39607, 39835, 40127, 40506, 40560, 40881, 41017, 41423, 41652, 41715, 42317, 42631, 42651, 42673, 42814, 43410, 43492, 43507, 43528, 43593, 44424, 44538, 44596, 44868, 45151, 45788, 46033, 46408, 46556, 46804, 46843, 46853, 46896, 46952, 47256, 47274, 47287, 47430, 47576, 47699, 47799, 48059,  
35 48125, 48227, 48359, 48378. The average fragment length was about 160 nt. The

observed sites of cleavage were largely (>95%) compatible with that predicted from an *MSE1/RcoR1* digest map.

5

### EXAMPLE 17

#### SINGLE NUCLEOTIDE EXTENSION ASSAYS

RNA preparation: Total RNA was isolated was prepared from Jurkat cells using (starting with  $1 \times 10^9$  cells in exponential growth) using an RNA isolation  
10 kit from Promega (WI). RNA was stored in two aliquots: 1) stock aliquote in diethyl pyrocarbonate-treated ddH<sub>2</sub>O was stored at -20°C, and 2) long term storage as a suspension in 100% H<sub>2</sub>O.

Reverse Transcription: Poly(dT) primed reverse transcription of total RNA was performed as described as described in Ausubel et al. (Ausubel et al., in  
15 Current Protocols in Molecular Biology, 1991, Greene Publishing Associates/Wiley-Interscience, NY, NY.) except that the reaction(s) were scaled to using 1 µg of input total RNA. 20-50 units of reverse transcriptase (Promega) was diluted 10-fold in 10% glycerol, 10 mM KPO<sub>4</sub> pH 7.4, 0.2% Triton X-100, and 2 mM DTT and placed on ice for 30 minutes prior to addition to the reactions. Gene-specific reverse transcription for  
20 GADPH and other control genes as described below were performed using 1 µg of total Jurkat RNA reversed transcribed in 10 mM Tris-HCl pH 8.3, 50 mM KCl, M MgCl<sub>2</sub>, 1 mM dNTPs, 2 U/µl RNasin (Gibco-BRL), 0.1 µM oligomer and 0.125 U/µl of M-MLV reverse transcriptase (Gibco-BRL) in 20µl reactions. Reactions were incubated in at 42°C for 15 minutes, heat inactivated at 95°C for 5 minutes , and diluted  
25 to 100 µl with a master mix of (10 mM Tris HCl pH 8.3, 1 mM NH<sub>4</sub>Cl, 1.5 M MgCl<sub>2</sub>, 100 mM KCl), 0.125 mM NTPs, 10 ng/ml of the respective oligonucleotide primers and 0.75 units of TAQ polymerase (Gibco-BRL) in preparation for PCR amplification.

PCR: PCR for each gene was performed with gene specific primers spanning a known intron/exon boundry (see below). All PCRs were done in 20µl  
30 volumes containing 10 mM Tris HCl pH 8.3, 1 mM NH<sub>4</sub>Cl, 1.5 M MgCl<sub>2</sub>, 100 mM KCl), 0.125 mM NTPs, 10 ng/ml of the respective oligonucleotide primers and 0.75 units of TAQ polymerase (Gibco-BRL). Cycling parameters were 94°C preheating step for 5 minutes followed by 94°C denaturing step for 1 minute, 55°C annealing step for 2 minutes, and a 72°C extension step for 30 seconds to 1 minute and a final extension at  
35 72°C for 10 minutes. Amplifications were generally 30-45 in number.

Purification of templates: PCR products were gel purified as described by Zhen and Swank (Zhen and Swank, *BioTechniques*, 14(6):894-898, 1993). PCR products were resolved on 1% agarose gels run in 0.04 M Tris-acetate, 0.001 M EDTA (1x TEA) buffer and stained with ethidium bromide while visualizing with a UV light source. A trough was cut just in front of the band of interest and filled with 50-200  $\mu$ l of 10% PEG in 1x TAE buffer. Electrophoresis was continued until the band had completely entered the trough. The contents was then removed and extracted with phenol, chloroform extracted, and then precipitated in 0.1 volume of 7.5 M ammonium acetate and 2.5 volumes of 100% EtOH. Samples were washed with 75% EtOH and briefly dried at ambient temperature. Quantitation of yield was done by electrophoresis of a small aliquot on 1% agarose gel in 1x TBE buffer with ethidium bromide staining and comparison to a known standard.

Each SNUPE reaction was carried out in a 50  $\mu$ l volume containing about 100 ng of the amplified DNA fragment, 1  $\mu$ M of the SNUPE primer, 2 units of Tag polymerase, and 1  $\mu$ l of the appropriate dNTP. All dNTPs are unlabelled in this type of assay. The buffer used was 10 mM Tris-HCl (pH 8.3), with 50 mM KCl, 5mM MgCl<sub>2</sub> and 0.001% (wt/vol) gelatin. The samples were subjected to one cycle consisting of a 2-minute denaturation period at 95°C, a 2 minute annealing period at 60°C and a 2-minute primer extension period at 72°C. The sequence of the SNUPE primer for each family is described below.

Primer extensions: Single nucleotide primer extensions were performed as described in Singer-Sam et al., (Singer-Sam et al., *PCR Methods and Applications* 1:160-163, 1992) except that 1 mM Mg<sup>++</sup>, 0.1  $\mu$ M primer, and 0.05  $\mu$ M of each dNTP type was used in each reaction type. After each primer extension described above, one-fifth volume of a loading dye (80% formamide, 0.1 % bromophenol blue, 0.1 % xylene cyanol, 2 mM EDTA) was added, and the entire sample electrophoresed in 15% denaturing polyacrylamide gel. Gels were fixed in 10% glycerol, 10% methanol, 10% glacial acetic acid with constant shaking followed by washing steps with 10% glycerol. The gels were then dried at 55°C for 3-5 hours.

The primers described in this experiment are described by Rychlik (Rychlik, *BioTechniques* 18:84-90, 1995) Primers may be synthesized or obtained as gel-flirtation grade primers from Midland Certified Reagent Company (Midland Texas). The amplifications are either TAQ DNA polymerase-based (10 mM Tris-HCl pH 8.3, 1.5 mM MgCl<sub>2</sub>, 50 mM KCl) or Pfu DNA polymerase-based based (20 mM Tris-HCl pH 8.3, 2.0 mM MgCl<sub>2</sub>, 10 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1% Triton X-100, 0.1

mg/ml bovine serum albumin). The total nucleoside triphosphate (NTPs) concentration in the reactions is 0.8 mM, the primer concentration is 200 nM (unless otherwise stated) and the template amount is 0.25 ng of bacteriophage lambda DNA per 20 µl reaction. Cycling parameters were 94°C preheating step for 5 minutes followed by 94°C  
 5 denaturing step for 1 minute, 55°C annealing step for 2 minutes, and a 72°C extension step for 30 seconds to 1 minute and a final extension at 72°C for 10 minutes. Amplifications were generally 30-45 in number.

Two regions in the bacteriophage lambda genome (GenBank Accession #J02459) were chosen as the priming sites for amplification. The 5'-primer has a stable  
 10 GC-rich 3'-end: the 3' primer is chosen so that a 381 bp product will result. The 5' forward primer is H17: 5'-GAACGAAAACCCCCCGC (SEQ ID No. \_\_\_\_). The 3'-reverse primer is RP17: 5'-GATCGCCCCCAAACACATA (SEQ ID No. \_\_\_\_).

The amplified product was then tested for the presence of a polymorphism at position 31245. The following primer was used in four single  
 15 nucleotide extension assays; SNE17: 5'-GAACGAAAACCCCCCGC (SEQ ID No. \_\_\_\_). The four single nucleotide extension assays were then carried as described above. All the reactions are then pooled and 5 µl of the pooled material was injected onto the HPLC column (SeraSep, San Jose, CA) without further purification.

HPLC was carried out using automated HPLC instrumentation (Rainin,  
 20 Emeryville, CA., or, Hewlett Packard, Palo Alto, CA). Unpurified SNEA products which had been denatured for 3 minutes at 95 prior into injection into an HPLC were eluted with linear acetonitrile (ACN, J.T. Baker, NJ) gradient of 1.8%/minute at a flow rate of 0.9 ml/minute, The start and end points were adjusted according to the size of the SNEA product. The temperature required for the successful resolution of the SNEA  
 25 molecules was 50°C. The effluent from the HPLC was then directed into a mass spectrometer (Hewlett Packard, Palo Alto, CA) for the detection of tags, with the results set forth in TABLE 2.

TABLE 2

Tagged Primer	ddNTP type	retention time	extended?
SNE17-487	ddATP	2.5 minutes	no
SNE17-496	ddGTP	2.5 minutes	no
SNE17-503	ddCTP	4.6 minutes	yes
SNE17-555	ddTTP	2.5 minutes	no

The results therefore indicate that the mass spectrometer tag (CMST) tag was detected at a retention time of 4.6 minutes indicating that the SNE17 primer was extended by one base (ddCTP) and therefore the polymorphism was position 31245 was in this case a "G". The SNE17-487, SNE17-496, and SNE17-555 tagged primers were not extended and their retention times on the HPLC was 2.5 minutes respectively.

### EXAMPLE 18

In this Example (18), all reactions were conducted in foil-covered flasks. The sequence of reactions A→F described in this Example is illustrated in Figures 19A and 19B. Compound numbers as set forth in this Example refer to the compounds of the same number in Figures 19A and 19B.

A. To a solution of ANP linker (compound 1, 11.2 mmol) and diisopropylethylamine (22.4 mmol) in  $\text{CHCl}_3$  (60 ml) was added allyl bromide (22.4 mmol). The reaction mixture was refluxed for 3 hours, stirred at room temperature for 18 hours, diluted with  $\text{CHCl}_3$  (200 ml), and washed with 1.0 M HCl (2 x 150 ml) and  $\text{H}_2\text{O}$  (2 x 150 ml). The organic extracts were dried ( $\text{MgSO}_4$ ) and the solvent evaporated to give compound 2 as a yellow solid.

To a mixture of compound 2 in  $\text{CH}_2\text{Cl}_2$  (70 ml), tris (2-aminoethyl) amine (50 ml) was added and the reaction mixture stirred at room temperature for 18 hours. The reaction was diluted with  $\text{CH}_2\text{Cl}_2$  (150 ml) and washed with pH 6.0 phosphate buffer (2 x 150 ml). The organic extracts were dried ( $\text{MgSO}_4$ ) and the solvent evaporated. The residue was subjected to column chromatography (hexane/EtOAc) to give 1.63 g (58%) of compound 3:  $^1\text{H}$  NMR ( $\text{DMSO}-d_6$ ):  $\delta$  7.85 (dd, 2H), 7.70 (t, 1H), 7.43 (t, 1H), 5.85 (m, 1H), 5.20 (q, 2H), 4.58 (q, 1H), 4.50 (d, 2H), 2.70 (m, 2H), 2.20 (br s, 2H).

B. To a solution of Boc-5-aminopentanoic acid (1.09 mmol) and NMM (3.27 mmol) in dry DMF (6 ml), was added HATU (1.14 mmol) and the reaction mixture stirred at room temperature for 0.5 hours. A solution of compound 3 (1.20 mmol) in dry DMF (1 ml) was added and the reaction mixture stirred at room temperature for 18 hours. The reaction was diluted with EtOAc (50 ml) and washed with 1.0 M HCl (2 x 50 ml) and D.I.  $\text{H}_2\text{O}$  (2 x 50 ml). The organic extracts were dried ( $\text{MgSO}_4$ ) and evaporated to dryness. The residue was subjected to column chromatography to give 420 mg (91%) of compound 4:  $^1\text{H}$  NMR ( $\text{DMSO}-d_6$ ):  $\delta$  8.65 (d, 1H), 7.88 (d, 1H), 7.65 (m, 2H), 7.48 (t, 1H), 6.73 (br s, 1H), 5.85 (m, 1H), 5.55 (m,



1H), 5.23 (q, 2H), 4.55 (d, 2H), 2.80 (m, 2H), 2.05 (t, 2H), 1.33 (s, 9H), 1.20-1.30 (m, 4H).

C. A solution of compound **4** (0.9 mmol) in HCl•1,4-dioxane (20 mmol) was stirred at room temperature for 2 hours. The reaction mixture was concentrated, dissolved in MeOH and toluene, and concentrated again (5 x 5ml) to give 398 mg (quantitative) of the compound **5**: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8.75 (d, 1H), 7.88 (d, 1H), 7.65 (m, 2H), 7.51 (t, 1H), 7.22 (m, 2H), 5.85 (m, 1H), 5.57 (m, 1H), 5.23 (q, 2H), 4.55 (d, 2H), 2.80 (m, 2H), 2.71 (m, 2H), 2.07 (s, 2H), 1.40-1.48 (br s, 4 H).

D. To a solution of compound **21** (0.48 mmol, prepared according to Example 20) and NMM (1.44 mmol) in dry DMF (3 ml), was added HATU (0.50 mmol) and the reaction mixture stirred at room temperature for 0.5 hours. A solution of compound **5** (0.51 mmol) in dry DMF (3 ml) was added and the reaction stirred at room temperature for 18 hours. The reaction mixture was diluted with EtOAc (75 ml) and washed with 5% Na<sub>2</sub>CO<sub>3</sub> (3 x 50 ml). The organic extracts were dried (MgSO<sub>4</sub>) and the solvent evaporated to give 281 mg (78 %) of compound **6**: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8.65 (d, 1H), 8.17 (d, 1H), 7.82-7.95 (m, 4H), 7.68 (m, 3H), 7.50 (t, 1H), 6.92 (d, 1H), 5.85 (m, 1H), 5.57 (m, 1H), 5.20 (q, 2H), 4.55 (d, 2H), 4.30 (q, 1H), 4.05 (q, 2H), 2.95 (m, 4H), 2.80 (m, 2H), 2.72 (m, 2H), 2.05 (s, 3H), 2.01 (t, 2H), 1.58-1.77 (m, 3H), 1.50 (m, 4H), 1.30 (q, 3H), 1.17-1.40 (m, 9H).

E. To a mixture of compound **6** (0.36 mmol) in THF (4 ml), was added 1 M NaOH (1 mmol) and the reaction stirred at room temperature for 2 hours. The reaction mixture was acidified to pH 7.0 with 1.0 M HCl (1 ml) and the solvent evaporated to give compound **7** (quantitative): <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8.65 (d, 1H), 8.17 (d, 1H), 7.82-7.95 (m, 4H), 7.68 (m, 3H), 7.50 (t, 1H), 6.92 (d, 1H), 5.52 (m, 1H), 4.30 (q, 1H), 4.05 (q, 2H), 2.95 (m, 4H), 2.80 (m, 2H), 2.72 (m, 2H), 2.05 (s, 3H), 2.01 (t, 2H), 1.58-1.77 (m, 3H), 1.50 (m, 4H), 1.30 (q, 3H), 1.17-1.40 (m, 9H).

F. To a solution of compound **7** (0.04 mmol) and NMM (0.12 mmol) in dry DMF (0.4 ml), was added HATU (0.044 mmol) and the reaction stirred at room temperature for 0.5 hours. Allylamine (0.12 mmol) was added and the reaction mixture stirred at room temperature for 5 hours. The reaction mixture was diluted with EtOAc (15 ml) and washed with 5% Na<sub>2</sub>CO<sub>3</sub> (3 x 10 ml). The organic extracts were dried (MgSO<sub>4</sub>) and the solvent evaporated to yield 15 mg (49%) of compound **8**: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8.49 (d, 1H), 8.17 (d, 1H), 7.82-7.95 (m, 4H), 7.68 (m, 3H), 7.50 (t, 1H), 6.92 (d, 1H), 5.72 (m, 1H), 5.50 (m, 1H), 5.03 (q, 2H), 4.37 (d, 2H), 4.30 (q, 1H),

4.05 (q, 2H), 2.95 (m, 4H), 2.80 (m, 2H), 2.72 (m, 2H), 2.05 (s, 3H), 2.01 (t, 2H), 1.58-1.77 (m, 3H), 1.50 (m, 4H), 1.30 (q, 3H), 1.17-1.40 (m, 9H).

### EXAMPLE 19

5

The sequence of reactions A→G as described in this Example 19 is illustrated in Figures 20A and 20B. Compound numbers as set forth in this Example refer to the compounds of the same number in Figures 20A and 20B.

A. To a solution of Fmoc-Lys(Boc)-OH (compound **9**, 33.8 mmol) in CHCl<sub>3</sub> (200 ml), was added diisopropylethylamine (67.5 mmol) and allyl bromide (67.5 mmol). The reaction mixture was refluxed for 6 hours, stirred at room temperature for 16 hours, diluted with CHCl<sub>3</sub>, washed with 1.0 M HCl (2 x 150 ml), saturated NaHCO<sub>3</sub> (1 x 150 ml) and D.I. H<sub>2</sub>O (2 x 150 ml). The organic extracts were dried (MgSO<sub>4</sub>) and the solvent evaporated to yield compound **10**.

15 To a solution of compound **10** in CHCl<sub>3</sub> (90 ml), was added pyrrolidine (10 eq.) and the reaction was stirred at room temperature for 2.5 hours. The reaction mixture was diluted with CHCl<sub>3</sub> (150 ml) and washed with saturated NaHCO<sub>3</sub> (3 x 250 ml). The organic extracts were dried (MgSO<sub>4</sub>) and the solvent evaporated. The residue was subjected to column chromatography (EtOAc / MeOH) to give 6.52 g (67%) of compound **11**: <sup>1</sup>H NMR (CDCl<sub>3</sub>): δ 5.90 (m, 1H), 5.27 (m, 2H), 4.60 (d, 2H), 3.48 (t, 1H), 3.10 (d, 2H), 1.40-1.78 (m, 9H), 1.40 (s, 9H).

B. To a solution of N-methylisonipecotic acid (1.60 mmol) and N-methyl morpholine (4.80 mmol) in dry DMF (5 ml), was added HATU (1.67 mmol). After 0.5 hours, a solution of compound **11** (1.75 mmol) in dry DMF (2 ml) was added and the reaction mixture stirred at room temperature for 18 hours. The reaction mixture was diluted with CH<sub>2</sub>Cl<sub>2</sub> (60 ml) and washed with saturated Na<sub>2</sub>CO<sub>3</sub> (3 x 40 ml). The organic extracts were dried (MgSO<sub>4</sub>) and the solvent evaporated. The residue was subjected to column chromatography (CH<sub>2</sub>Cl<sub>2</sub> / MeOH / triethylamine) to give 580 mg (88%) of compound **12**: <sup>1</sup>H NMR (DMSO): δ 8.12 (d, 1H), 6.77 (t, 1H), 5.90 (m, 1H), 5.27 (m, 2H), 4.53 (d, 2H), 4.18 (m, 1H), 2.62-2.90 (m, 5H), 2.13 (s, 3H), 1.85 (m, 2H), 1.57 (m, 5H), 1.35 (s, 9H), 1.00 (t, 2H).

C. A mixture of compound **12** (1.39 mmol) in HCl•1, 4-dioxane (20 mmol) was stirred at room temperature for 4 hours. The reaction mixture was concentrated, dissolved in MeOH, coevaporated with toluene (5 x 5ml) to give 527 mg (quantitative) of compound **13**: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8.12 (d, 1H), 6.77 (t, 1H), 5.90

35

(m, 1H), 5.27 (m, 2H), 4.53 (d, 2H), 4.18 (m, 1H), 2.65-3.00 (m, 8H), 2.23 (s, 3H), 1.85 (m, 2H), 1.57 (m, 5H), 1.00 (t, 2H).

D. To a solution of 4-ethoxybenzoic acid (1 eq.) in dry DMF, is added NMM (3 eq.) and HATU (1.05 eq.). After 0.5 hours, a solution of compound **13** in dry DMF is added. After the completion of the reaction and basic workup, the compound **14** is isolated and purified.

E. To a solution of compound **14** in THF, is added 1N NaOH and the reaction mixture stirred at room temperature. After the completion of the reaction and acidification, the compound **15** is isolated.

F. To a solution of compound **15** (1 eq.) in dry DMF, is added NMM (3 eq.) and HATU (1.05 eq.). After 0.5 hours, a solution of compound **21** (ANP - allyl ester, prepared according to Example 20) in dry DMF is added. After the completion of the reaction and basic workup, the title compound **16** is isolated and purified.

G. To a solution of compound **16** in THF, is added 1N NaOH and the reaction mixture stirred at room temperature. After the completion of the reaction and acidification, the compound **17** is isolated.

#### EXAMPLE 20

The sequence of reaction A through D as described in this Example 20 is illustrated in Figure 21. Compound numbers as set forth in this Example, as well as Examples 18 and 19, refer to the compounds of the same number in Figure 21.

A. To a solution of 4-ethoxybenzoic acid (7.82 mmol) and N-methyl morpholine (20.4 mmol) in  $\text{CH}_2\text{Cl}_2$  (10 ml), was added HATU (7.14 mmol). After 0.25 hours, a solution of compound **11** (6.8 mmol) in  $\text{CH}_2\text{Cl}_2$  (6 ml) was added and the reaction mixture stirred at room temperature for 18 hours. The reaction was diluted with  $\text{CH}_2\text{Cl}_2$  (150 ml) and washed with 1.0 M HCl (3 x 50 ml) and saturated  $\text{NaHCO}_3$  (3 x 50 ml). The organic extracts were dried ( $\text{MgSO}_4$ ) and the solvent evaporated. The residue was subjected to column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$ ) to give 2.42 g (82%) of compound **18**:  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  7.78 (d, 2H), 6.91 (d, 2H), 6.88 (d, 1H), 5.83-5.98 (m, 1H), 5.21-5.38 (m, 2H), 4.80 (q, 1H), 4.66 (d, 2H), 4.06 (q, 2H), 3.11 (q, 2H), 1.90-2.04 (m, 1H), 1.68-1.87 (m, 1H), 1.39 (t, 3H), 1.34 (s, 9H), 1.32-1.58 (m, 4H).

B. A mixture of compound **18** (5.5 mmol) in HCl • 1, 4-dioxane (14.3 mmol) was stirred at room temperature for 1 hour. The reaction mixture was

concentrated, dissolved in MeOH, azeotroped with toluene, and concentrated again (5 x 5 ml) to give a quantitative yield of compound **19**.

C. To a solution of N-methylisonipecotic acid (6.21 mmol) in dry DMF (15 mL), was added NMM (21.6 mmol) and HATU (5.67 mmol). After 0.5 hours, a solution of compound **19** (5.4 mmol) in dry DMF (10 ml) was added and the reaction stirred at room temperature for 18 hours. The reaction mixture was brought to pH 12 with 1N NaOH (20 ml) and extracted with CHCl<sub>3</sub> (2 x 200 ml). The organic extracts were dried (MgSO<sub>4</sub>) and the solvent evaporated to give 2.2 g (89%) of compound **20**: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 8.52 (d, 1H), 7.84 (d, 2H), 7.72 (t, 1H), 6.95 (d, 2H), 5.80-5.95 (m, 1H), 5.18-5.31 (dd, 2H), 4.58 (d, 2H), 4.37 (q, 1H), 4.08 (q, 2H), 3.01 (d, 2H), 2.08 (s, 3H), 1.95 (m, 1H), 1.63-1.82 (m, 4H), 1.51 (m, 4H), 1.32 (t, 3H), 1.22-1.41 (m, 6H).

D. To a solution of compound **20** (4.4 mmol) in THF (10 ml), is added 1N NaOH (4.4 mmol) and the reaction mixture stirred at room temperature for 1 hour. The reaction was concentrated, dissolved in THF/toluene (2 x 5 ml), concentrated, dissolved in CH<sub>2</sub>Cl<sub>2</sub>/toluene (1 x 5 ml) and concentrated again to give a quantitative yield of compound **21**: <sup>1</sup>H NMR (DMSO-d<sub>6</sub>): δ 7.76 (d, 2H), 6.96 (d, 2H), 4.04 (q, 2H), 3.97 (d, 1H), 2.97 (d, 2H), 2.64 (d, 2H), 2.08 (s, 3H), 1.95 (m, 1H), 1.58-1.79 (m, 4H), 1.44 (m, 6H), 1.30 (t, 3H), 1.11-1.35 (m, 4H).

### EXAMPLE 21

The synthesis of the CMSTs (Cleavable, Mass Spectrometry-detectable Tags) may be based on a combinatorial approach as described in Figure 22. The general approach is designed to be compatible with developments of the mass spectrometers and changes and improvements in ionization technologies. A central scaffold is first tested for compatibility with the type of ionization that is to be employed in the method. It is important that the scaffold not be susceptible to fragmentation, heat degradation, or dimerization to adduct formation. With the current APCI/quadrupole mass spectrometers, about 400 tags will fill the spectrum taking into account isotopic contamination which forces a minimal spacing of about 4 AMU.

The CMSTs were designed in a modular fashion such that the tags can be built with a combinatorial chemistry approach. There are 5 "modules" (functionally separate atomic groups), in the complete tagged-oligonucleotide. The first module is the oligonucleotide (ODN), which can be of any length and sequence and which

possesses a 5'-hexylamine for coupling to the linker, which is preferably a photocleavable linker. The ODN can serve either as a probe for hybridization or a primer in an extension, ligation, or enzymatic-based assay. The 2nd module is the linker, preferably a photo-labile linker, that connects the tag to the oligonucleotide probe or primer. In the approach described in detail here, the photolabile linker is a o-nitrobenzyl amino acid derivative (see Lebl M. et al. "Dynamic database of references in molecular diversity" having internet address <http://www.5z.com>. for a compendium). The 3rd module is the ionization enhancer. This module is the scaffold upon which the CMST is synthesized, and in general provides functionally that is specific for the type of ionization method used (i.e., APCI (positive or negative mode), electrospray, MALDI, etc.). The 4th module is the gross mass adjuster upon which the mass can be altered in large increments of 200-500 amu. This module allows the re-use of the variable weight adjusters. The fifth module is the variable weight adjuster, also referred to variable mass units (VMU). Variable weight adjusters are chemical sidearms which are added to the tag scaffold. These variable weight adjusters fine-tune the weight of the CMSTs. The weight of the tags is spaced at least every 4 amu to avoid overlapping spectra due to isotopic contamination. The same VMU sidearms may be repeatedly used with the different types of ionization scaffolds. In summary, the ionization module and variable mass adjusters are designed to confer predictable behaviour in the MSD, the photocleavage of the photolabile linker is fast (described below), the CMST/oligonucleotide conjugates are compatible with with PCR and HPLC and other manipulations found in assay formats.

A detailed synthetic route of the tags is described below and in Figure 22. The synthetic route towards a CMST begins with the esterification of the photosensitive ANP linker (3-amino-3-(2-nitrophenyl) propionic acid, (1) to give the ethyl ester hydrochloride (2) in 84% yield. An important step in the process is the enzymatic transformation of (2) to provide the ethyl ester as a single isomer. After the ethyl ester hydrochloride has been basified to the free amine and concentrated, the oily residue was brought up in pH 7 phosphate buffer and adjusted to neutral pH with 2N HCl. The Amano PS esterase enzyme was added as a phosphate buffer slurry. After the completion of the reaction, a basic workup removed the hydrolyzed ANP by-product (4), and the single-isomer ethyl ester (3, >99% e.e.) was recovered (92% of available material).

Coupling of (3) with a-BOC-e-alloc-lysine (5), using EDAC and HOBT, gave the protected ANP lysine (6) in 91% yield. Removal of BOC with TFA provided

the amino-e-alloc-lysine ANP ester (7) as a white solid. Methyl isonipecotic acid hydrochloride was coupled to 7, using EDAC and triethylamine, to give crude alloc-protected core structure (8), which was deprotected with diethylamine, triphenylphosphine, and palladium acetate, at 50 degrees C. The resulting core structure (9) was crystallized from the reaction mixture and recovered by filtration in 95% yield as a yellow solid.

A variety of carboxylic acids, designated variable mass units (VMU), were coupled to (9), using HATU and N-methyl morpholine. A set of VMU's were designed to provide even-mass tags and 4 a.m.u. spacing, to avoid isotopic contamination. The following were used as exclusion criteria in selecting particular VMU's at the target masses: 1) functionalities incompatible with the synthetic sequence (e.g. esters); 2) elements with multiple isotopes (e.g., Cl, Br, S); 3) functionalities that might lead to competing photoprocesses (iodides, acyl- and aryl-phenones); 4) racemic acids; and 5) availability from vendors.

After purification by column chromatography, the CMST ethyl ester (10) was recovered in variable yields. Base hydrolysis of (10) with NaOH gave the CMST acid (11) in quantitative yield. The final step, the formation of the activated ester, was achieved using tetrafluorophenol trifluoroacetate and Hunnrich's base, and resulted in the CMST TFP ester (12) in variable yields.

It is convenient to attach the tags (in general) to the 5'-end of oligonucleotides so the 3' hydroxyl can be extended in the polymerase chain reaction or be available for other enzymatic modifications. Also, when used directly as probes as described in the application here, the tags are preferably attached to the 5'-end of the oligonucleotide probes. The guidelines described by Lukhtanov et. al., "Oligodeoxyribonucleotides with conjugated dihydropyrroloindole oligopeptides: preparation and hybridization properties", *Bioconjug Chem* 6(4):418-26, Jul-Aug, 1995, may be followed to prepared the tagged oligonucleotides from CMSTs and oligonucleotides.

## EXAMPLE 22

A PHRED photocleavage unit is placed between the HPLC and the mass spectrometer or an autoinjector and the MSD. "PHRED" which stands for Photochemical Reactor for Enhanced Detection, and is available from Aura Industries, Staten Island, NY (available with both a 254 nm and 366 nm bulb, the 254 nm bulb was

used). An inline device is preferably placed between the separation instrumentation (e.g., HPLC or gel) and the detector. The interface preferably has the following properties: the ability to collect the DNA fragments at discreet time intervals, concentrate the DNA fragments, remove the DNA fragments from the electrophoresis buffers and milieu, cleave the MW-identifier from the DNA fragment, separate the MW-identifier from the DNA fragment, dispose of the DNA fragment, place the tag in a volatile solution, volatilize and ionize the tag which introduces the tag into mass spectrometer.

A suitable configuration of the photocleavage device is a 300 cm long, 8 watt, UV germicidal lamp (G8T50) (with an emission around 366 nm) under which a 80 l coil of 0.01 inch ID tefzel tubing is placed. Flow rates of 800  $\mu$ l per minute are generally suitable. Solution compositions which are compatible with APCI/MS and which contain only low concentrations of acetonitrile and buffers such as Tris-HCl are preferred. With this configuration, there is no requirement to separate the DNA from the tag prior to the ionization step. The photolabile linker cleaves very rapidly under these conditions. The heat source in the APCI chamber contributes to the cleavage of the photolabile linker.

By varying the length of the tefzel tubing coil under the UV source and holding the rate of flow constant, the residence time under the UV source was varied from the 0.75 to 6 seconds. The response factor (the integrated ion current produced per mole tag injected into the flow stream in "flow injection analysis" (FIA)) of the tags was determined for a pool of 6 tags all tethered to a single oligonucleotide sequence (a 20-mer, M13 forward sequencing primer). The response factor is the integral of the efficiency of ionization, ion introduction into the vacuum chamber, and then subsequent detection by the MSD. Each tag in the pool was present at a concentration of 100 fmol per  $\mu$ l. The diluent was tRNA (Boehringer Mannheim) at a concentration of 1  $\mu$ g/ml in HPLC-grade water. In TABLE 3, the response factor and the percentage of the signal observed relative to the longest exposure (6 seconds) are listed on a per tag basis. The results indicate that the tag is rapidly cleaved (less than 2 seconds) from the tethered oligonucleotide. In the time frame of 1.9 to 6 seconds there was little decrease in the observed response factor with 6 different tagged-ODNs. At the shortest exposure time tested (0.75 seconds) there was up to a 25% decrease in the observed response factor.

TABLE 3

Tag MW	10 $\mu$ l	25 $\mu$ l	50 $\mu$ l	80 $\mu$ l
--------	------------	------------	------------	------------

447	4300 81%	4400 83%	5300 100%	5300 100%
455	4900 92%	5200 98%	6400 120%	5300 100%
479	3600 75%	4100 85%	4900 102%	4800 100%
503	4600 74%	5700 92%	6700 108%	6200 100%
507	5200 85%	5400 88%	6500 104%	6100 100%
511	4100 78%	5100 98%	5500 105%	5200 100%

There was an approximate 40% decrease in the response factor of the six tags when the lamp in the photocleavage unit (data not shown). It has been determined that the photolabile linker is thermolabile and the tag is apparently cleaved during the APC ionization step where the vaporization temperature is 450°C.

### EXAMPLE 23

#### COLLECTIVE TAG BEHAVIOR AND RELATIVE STABILITY

Tags were designed to provide a single parental ion that did not fragment into daughter ions, or form adducts. The response of the tags was determined for a pool of 43 tags all tethered to a single oligonucleotide sequence (20, M13 sequence). Each tag in the pool was present at a concentration of 100 fmol per  $\mu$ l (the concentration was determined by applying a dilution factor to a oligo/tag stock that was measured spectrometrically at 260 nM. A correction for the tag contribution to absorbance at 260 nM was not used. The diluent was tRNA (Boehringer Mannheim) at a concentration of 1  $\mu$ g/ml in HPLC-grade water. The pool was stored at 4°C in the dark. CMST-tagged oligonucleotides may be handled under normal laboratory lighting conditions without occurrence appreciable degradation.

Prior to analysis, a 55  $\mu$ l aliquot was removed from the stock solution and placed in a 200  $\mu$ l polypropylene autosampler vial and crimped closed. Five injections of 5  $\mu$ l each from the pool and 3 injections of 5  $\mu$ l each from the diluent was preformed by the HPLC HP1100 ALS. The APCI-MS chamber parameters were as follows: 20 PSI for the nebulizer pressure, vaporizer temperature was 450°C, drying gas flow was 3 L/min, drying gas temperature was 350°C, corona current was 4 uA, fragmentor voltage 125 V, the gain was set to 1, and the peak width 0.07 minutes. The flow rate was 0.8 ml/min., the "dead space" of the photocleavage unit was 80  $\mu$ l, (0.01" ID Tefzel). The lamp of the photocleavage unit operated at 366 nM. Each tag was



quantified by extraction of the SIM ion from the TIC. Peaks were integrated under the following parameters: slope sensitivity at 2500, minimum peak area at 800, minimum peak height at 100, peak width at 0.15, shoulders settings at "off". Peak area for each tag was recorded for all five injections within an experiment and the average area was calculated along with the standard deviation and coefficient of variation. Average areas, standard deviations, and CVs were also calculated for a single day, over a two day period and for a three day period. Within a single day the coefficient of variation varied between 2.0% and 9.9% between different tags. Over a three day period the coefficient of variation varied between 4.0% and 9.8% between different tags. The tags are therefore stable with respect to storage and pooling, as shown in TABLES 4a, 4b and 4c.

**TABLE 4a**  
**1-DAY**

Tag MW	CV @ 10 fmol/injection	CV @ 100fmol/injection
447	2.0	3.7
455	4.0	6.5
479	6.9	7.9
503	9.9	2.9
507	4.1	1.8
511	4.1	4.0

**TABLE 4b**  
**2-DAY**

Tag MW	CV @ 10 fmol/injection	CV @ 100fmol/injection
447	3.0	4.1
455	3.8	8.6
479	4.8	6.7
503	1.2	5.6
507	5.5	8.7
511	4.5	5.6

**TABLE 4c**  
**3-DAY**

Tag MW	CV @ 10 fmol/injection	CV @ 100fmol/injection
--------	------------------------	------------------------

447	3.3	4.9
455	4.0	3.9
479	5.6	6.7
503	9.8	8.9
507	2.0	4.8
511	3.8	6.7

### Sensitivity and Lower Limit of Detection

Current mass produced quadrupoles have sensitivities comparable to fluorescent based sequencing. The sensitivity can be expressed in terms of the lower level of detection of the mass of the CMST. We are defining the lower limit of detection as 3 standard deviations above the background of the assay system. The lower limit of detection of the pool of 43 conjugates was determined for the 30-ion SIM mode.

A set of 10, two-fold dilutions were prepared in the tRNA/water diluent to give 500, 250, 125, 62.5, 31.2, 15.6, 7.8, 3.9, 1.9 and 0.9 femptomole of material per injection (5-10  $\mu$ l). The data was obtained in the 30-ion SIM mode under conditions described above. The LLD for each tag is shown in TABLE 5.

TABLE 5

THE LOWER LIMIT OF DETECTION FOR 43 CMSTs.

Tag MW	LLD	Signal at LLD
367	30	1970
371	30	1707
375	15	1318
379	15	1642
383	30	2585
387	15	1301
391	15	1554
395	15	1960
403	8	1784
407	15	2266
411	15	2285

Tag MW	LLD	Signal at LLD
415	8	1362
419	8	1260
423	8	1530
427	15	1607
431	8	1606
439	8	2797
443	4	1576
447	4	1144
459	8	1881
463	8	1179
467	8	1625
471	2	1094
475	4	1664
479	2	2626
483	8	1528
487	8	1556
495	15	1894
499	15	1465
503	4	1208
507	8	1843
511	8	1453
515	8	1276
519	8	1042
523	15	1690
527	8	815
531	8	1521
535	8	1276
539	4	1651
543	15	1076

The LLD for this particular set of tags is approximately 4-30 femtomoles per injection following the photocleavage step from the oligonucleotide: (30 x 10<sup>-15</sup> moles @ 500 MW ---> 5 x 10<sup>-11</sup> gram ---> 50 picograms of tag, 100 x 10<sup>-15</sup> mole at

330 x 400 nt x 2 = 3 x 10<sup>-8</sup> gram = 30 nanograms of a 400 nt double strand PCR product). Therefore, assuming an average 25 µl PCR reaction contains 300 ng of a double strand PCR product, about 1/10th of that product will be used to generate a signal that is statistically above the background of the measurement. For a means of comparison, approximately 100-200 ng of DNA per lane on the ABI 377 sequencer was used.

The lower limit of detection of the tagging system effects the level of multiplexing that can be achieved in a single injection. To date, we have not noted any problem with the measurement process when up to 10 µg of PCR product is used per injection. The implication of this number is that 10,000/30 or about 300 reactions can be multiplexed using the tagging system of the present invention. Therefore, for the HP-MSD1100, the level of multiplexing is about the same as the tags that can be placed in the spectrum of the quadrapole MS.

#### 15 **Tag Interference**

There is little or no interference in terms of response (response being the measureable ion current that reflects the extent of ionization, degree of fragmentation or degradation, adduct formation, etc.) of multiplexing large numbers of CMSTs of the present invention. The multiplexing of tags does not affect the response of individual tags and therefore multiplexing does not affect ionization or total ion current.

#### **Response factor as a function of injection volume**

As described above, a pool of 43 conjugates were measured to determine the response factor as a function of injection volume. 50 fmol per injection and 500 fmol per injection were measured in 5, 10, 20, 50 and 100 µl volumes. The APCI-MS parameters were identical to those used in the tag stability experiment. The values shown in TABLE 4 are an average of 5 replicate injections for each volume measured. There was no decrease in the RF between 5 µl and 10 µl volumes. At the 20 µl injection volumes the RF was 90-97% of the 5 µl volume. When 50 µl volumes were measured, the RF was 54%-75% of the 5 µl volume. There was no detectable signal using the 100 µl injection volumes.

EXAMPLE 24

## P4502D6 POLYMORPHISM DETECTION

The CMST technology platform described herein may be used to  
5 measure the CYP2D6 polymorphisms which are responsible for the metabolism of  
debrisoquine 4-hydroxylase. This particular P450 cytochrome is important in the  
metabolism of more than 30 drugs and xenobiotic compounds. CYP2D6 (P450-2D6) is  
estimated to be responsible for metabolizing approximately 25% of the currently  
prescribed drugs. The CYP2D6 is also known as debrisoquine/sparteine hydroxylase.  
10 Sachse, *Am. J. Human Genetics*, 60:284-295, 1997, has estimated that up to 10% of  
Caucasians are poor metabolizers due to either inactivating mutations in both alleles or  
complete lack of the gene. There is also a broad range of CYP2D6 activities in addition  
to poor metabolizers. The mutations selected for study include C188T, G212A,  
delT1795, G1846T/A, G1934A, delA2637, C2938T, and G4268C. The RFLPS were  
15 detected by gel electrophoresis as previously described (see Gough et. al., *Nature*  
347:773-776, 1990). Primers used for RFLP corresponds to those used by Sachse et al,  
*Am. J. Hum. Genet.* 60:284-295, 1997. Primers used for sequencing corresponds to  
those used by Meyer et al, *Pharmacogenetics* 5:373-384, 1995.

The principle of the CMST-based assay was to immobilize one strand of  
20 the amplified CYP2D6 exon on a solid phase (e.g., magnetic particles), hybridize the  
oligonucleotide probes, wash away unhybridized material, elute the hybridized probe  
and then detect the mass spec tag by mass spectrometry (after cleaving the tag from the  
probe).

The amplification conditions were as follows. Primers flanking the 2D6  
25 gene (Sachse et. al., *Am. J. Hum. Genet.* 60:284-295, 1997) were used to amplify a  
4,681b.p. genomic DNA fragment containing all of the relevant gene sequence. The  
PCR reaction was composed of 1X Expand HF buffer, 1.5mM MgCl<sub>2</sub>, 200μM dNTP's,  
0.5μM primers P100 & P200, 0.5% formamide, 100ng gDNA, and 1.1U Expand<sup>TM</sup>  
High Fidelity enzyme mix (Boehringer Mannheim). Thermocycling conditions were as  
30 follows: 94°C for 3 minutes; 10 cycles of 94°C for 30 seconds, 62°C for 30 seconds,  
and 68°C for 4 minutes; 20 cycles of 94°C for 30 seconds, 62°C for 30 seconds, and  
68°C for 4 minutes +20 seconds/cycle; 68°C for 10 minutes. Product were visualized  
on a 1.0 % agarose gel stained with ethidium bromide.

The assay format is described as follows. Streptavidin magnetic  
35 particles (Promega Magnesphere, binding capacity of 80 pmol biotin/100 μg particles)

were washed with low salt wash and binding buffer (LSWBB, 100 mM NaCl, 1 mM EDTA, 10 mM Tris, pH 7.5) and then resuspended in high salt wash and binding buffer (HSWBB, 2 M NaCl, 1 mM EDTA, and 10 mM Tris pH 7.5) at a concentration of 2000 µg/ml. The biotinylated PCR products were incubated with the streptavidin particles for 2 hours at 21°C with constant rotary mixing. The particles were washed twice with 200 µl of HSWBB and once with 200 µl of LSWBB. The bound PCR amplicons were then denatured by treatment with 50 µl of 0.1 N NaOH for 10 minutes at 21°C. The particles were then washed once with 50 µl of 0.1 N NaOH and three times with 200 µl LSWBB. The particle-bound amplicons were then hybridized with equal molar mixtures of wild-type (wt) and mutant (mt) probes possessing different mass tags. Fifty picomoles of respective probe was placed in 200 µl of 2 M GuSCN, 5 mM EDTA and 10 mM Tris pH 7.5 and 50 µl of the hybridization solution was placed with the particles. Hybridization was for 1 hour at 21°C with constant rotary mixing. The particles were washed 5 times with LSWBB and the tubes were changed after the second wash. The hybridized probes were eluted from the particles by treatment of the particles with 20 µl of 0.1 N NaOH and a following wash of 9 µl of 0.1 N NaOH. The solution was then neutralized with 3 µl of 1 M acetic acid. Five µl of this solution was then injected into the mass spectrometer (HP 1100 series LC/MS equipped with a vacuum degasser, binary pump, autosampler and diode array detector). The mass spectrometer was used with the APCI source option. HP LC/MSD Chemstation software was used for system control, data acquisition and data analysis installed on a HP vectra XA with the Windows NT workstation version 4.0 operating system). The flow stream into the MS consists of 50% acetonitrile in ultra-pure water at a flow rate of 800 µl/minute. The photochemical cleavage device consisted of 254 nm low pressure mercury lamp, a UV transparent reactor coil and a lamp holder (Aura Industries).

Representative results are shown in TABLE 6.

TABLE 6

Individual	Exon	mAU (wt)	mAU (mt)	CMST call	RFLP call
1362 PF 13	4	0	190,000	M/M	M/M
1362 PM 14	4	152,000	0	W/W	W/W
1362 MF 15	4	149,000	53,000	W/M	W/M
1377 C1 19	6	0	271,000	M/M	M/M
1377 C2 20	6	104,000	88,000	W/M	W/M
1377 C3 21	6	290,000	0	W/W	W/W
1377 C1 19	9	0	74,000	M/M	M/M
1377 C2 20	9	38,000	41,000	W/M	W/M
1377 C3 21	9	149,000	0	W/W	W/W
CONTROL		0	0	NONE	NONE

EXAMPLE 25

## 5 GENE EXPRESSION MONITORING WITH CMST-TAGGED ODNs

Total RNA (1-2ug) from the A549 human cell line was transcribed using Superscript II reverse transcriptase and oligodT-primer in a final volume of 22μl according to manufacturer's instructions (Life Technologies; Gaithersburg, MD). A

10 75bp region of the apoptosis-related human DAD-1 gene coding region spanning an intro-exon boundary was amplified by Taq polymerase chain reaction (PCR); an initial denaturation of 95°C (5 min) was followed by 25-60 cycles (annealing at 45°C for 15 sec and denaturation at 95°C for 5 sec). The PCR reaction (20-200μl) contained 0.5μM DAD-1 reverse primer (5'-biotin-CCA GGA AAT TCA AAG AGT GA-3'), 0.0125μM

15 or 0.5μM DAD-1 phosphorylated forward primer (5'-TTG GCT GAA TCA TTC TCA TT-3'),  $7 \times 10^2$ - $1.2 \times 10^7$  molecules of internal standard (5'-CC AGG AAA TTC AAA GAG TGA ACA TTC TTT TTG TGT CG-3'), 1μl A549 cDNA or  $8 \times 10^4$ - $10^{11}$  molecules of WT mimic (5'-CCA GGA AAT TCA AAG AGT GAA CAT TCT TTT AGT CTC CTA CTC CTC AAT TAA GTA AAT GAG AAT GAT TCA GCC AA-3'),

20 0.8U Taq polymerase, 0.2mM dATP, 0.2mM dCTP, 0.2mM dGTP, 0.2mM dTTP, 1.5mM MgCl<sub>2</sub>, 50mM KCl, and 10mM Tris-HCl pH 8.3.

The amplification product (50μl) was rendered single-stranded by either assymetric amplification conditions (above) or by digestion with 2.5U Lambda exonuclease (Boehringer Mannheim; Indianapolis, IN) for 15 min (37°C) in 18mM

25 Tris-HCl pH 9.5, 1.8mM MgCl<sub>2</sub>, 28mM KCl. The digested amplicon or assymetric PCR

reaction was placed at 70°C for 5 min, cooled briefly (room temperature), and adjusted to 2.1M guanidine isothiocyanate, 50mM Tris-HCl pH 7.5, 0.1% sarkosyl, 1µg/ml tRNA, 5µg/ml WT 394 probe (5'-394 MW CMST tag-TTG AGG AGT AGG AGA CTA AAA-3'), 5µg/ml IS 390 probe (5'-390 MW CMST tag-TTGACGACTACGACACAAAAA-3').

The hybridization reaction was incubated 10 min (RT) and transferred to a 30K MWCO spin filter (Millipore Corp.; Bedford, MA) containing 7µl of 286µg/ml Avidin-DN (Vector; Burlingame, CA) and 0.7µg/ml tRNA. The spin filter was incubated 5-10min (RT) and centrifuged (4,000xg, 10min). The spin filter was washed twice with four-hundred microliters of cold HPLC-grade dH<sub>2</sub>O and centrifuged (4,000xg, 10min). Twenty-five microliters of 1µg/ml tRNA were added to the spin filter to elute the retained hybrid. The 25µl retentate was injected in 50% acetonitrile through a photolysis unit into an HP mass spectrometer (APCI positive mode). Single ion measurements were made for tags having molecular weights of 390 and 394, and the results displayed as peak areas. The ratio of the 390 / 394 tag signals was proportional to the amount of amplicon generated from either the internal standard or the A549 DAD-1 cDNA or WT mimic. Calculation of the unknown number of RNA molecules in the A549 total RNA or number of WT mimics could be made from a standard curve as shown on the log/log plot of the number of input internal standard molecules vs. the ratio of the 390 / 394 tag signals. The unknown is the value of x when y=0. Alternatively, the unknown = the ratio of 390 / 394 X the number of input internal standard molecules, when the ratio is between 0.3 and 3.

The results are shown in Figure 23.

25

### EXAMPLE 26

#### SINGLE NUCLEOTIDE EXTENSION ASSAY

RNA preparation: Total RNA was isolated was prepared from Jurkat cells using (starting with 1 x 10<sup>9</sup> cells in exponential growth) using an RNA isolation kit from Promega (WI). RNA was stored in two aliquots: (1) stock aliquots in diethyl pyrocarbonate-treated ddH<sub>2</sub>O was stored at -20°C, and (2) long term storage as a suspension in 100% H<sub>2</sub>O.

Reverse Transcription: Poly(dT) primed reverse transcription of total RNA was performed as described as described in Ausubel et al. (Ausubel et. al., in *Current Protocols in Molecular Biology*, 1991, Greene Publishing Associates/Wiley-



Interscience, NY, NY.) except that the reaction(s) were scaled to using 1 µg of input total RNA. 20-50 units of reverse transcriptase (Promega) was diluted 10-fold in 10% glycerol, 10 mM KPO<sub>4</sub> pH 7.4, 0.2% Triton X-100, and 2 mM DTT and placed on ice for 30 minutes prior to addition to the reactions. Gene-specific reverse transcription for  
5 GADPH and other control genes as described in the figures and tables were performed using 1 µg of total Jurkat RNA reversed transcribed in 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1 mM MgCl<sub>2</sub>, 1 mM dNTPs, 2 U/µl RNasin (Gibco-BRL), 0.1 µM oligomer and 0.125 U/µl of M-MLV reverse transcriptase (Gibco-BRL) in 20 µl reactions. Reactions were incubated in at 42°C for 15 minutes, heat inactivated at 95°C for 5 minutes, and  
10 diluted to 100 µl with a master mix of (10 mM Tris HCl pH 8.3, 1 mM NH<sub>4</sub>Cl, 1.5 M MgCl<sub>2</sub>, 100 mM KCl), 0.125 mM NTPs, 10 ng/ml of the respective oligonucleotide primers and 0.75 units of TAQ polymerase (Gibco-BRL) in preparation for PCR amplification.

PCR: PCR for each gene was performed with gene specific primers  
15 spanning a known intron/exon boundry (see tables below). All PCRs were done in 20 µl volumes containing 10 mM Tris HCl pH 8.3, 1 mM NH<sub>4</sub>Cl, 1.5 M MgCl<sub>2</sub>, 100 mM KCl), 0.125 mM NTPs, 10 ng/ml of the respective oligonucleotide primers and 0.75 units of TAQ polymerase (Gibco-BRL). Cycling parameters were 94°C preheating step for 5 minutes followed by 94°C denaturing step for 1 minute, 55°C annealing step for 2  
20 minutes, and a 72°C extension step for 30 seconds to 1 minute and a final extension at 72°C for 10 minutes. Amplifications were generally 30-45 in number.

Purification of templates: PCR products were gel purified as described by Zhen and Swank (Zhen and Swank, 1993, *BioTechniques*, 14(6), p894-898.). PCR products were resolved on 1% agarose gels run in 0.04 M Tris-acetate, 0.001 M EDTA  
25 (1x TEA) buffer and stained with ethidium bromide while visualizing with a UV light source. A trough was cut just in front of the band of interest and filled with 50-200 µl of 10% PEG in 1x TAE buffer. Electrophoresis was continued until the band had completely entered the trough. The contents was then removed and extracted with phenol, chloroform extracted, and then precipitated in 0.1 volume of 7.5 M ammonium  
30 acetate and 2.5 volumes of 100% EtOH. Samples were washed with 75% EtOH and briefly dried at ambient temperature. Quantitation of yield was done by electrophoresis of a small aliquot on 1% agarose gel in 1x TBE buffer with ethidium bromide staining and comparison to a known standard.

Each SNUPE reaction was carried out in a 50 µl volume containing about  
35 100 ng of the amplified DNA fragment, 1 µM of the SNUPE primer, 2 units of Tag

polymerase, and 1  $\mu$ l of the appropriate dNTP. All dNTPs are unlabelled in this type of assay. The buffer used was 10 mM Tris-HCl (pH 8.3), with 50 mM KCl, 5mM  $MgCl_2$  and 0.001% (wt/vol) gelatin. The samples were subjected to one cycle consisting of a 2-minute denaturation period at 95°C, a 2 minute annealing period at 60°C and a 2 minute primer extension period at 72°C. The sequence of the SNUPE primer for each family is described below.

Primer extensions: Single nucleotide primer extensions were performed as described in Singer-Sam et. al., (Singer-Sam et. al., 1992, *PCR Methods and Applications*, 1, p160-163) except that 1 mM  $Mg^{++}$ , 0.1  $\mu$ M primer, and 0.05  $\mu$ M of each dNTP type was used in each reaction type. After each primer extension described above, one-fifth volume of a loading dye (80% formamide, 0.1 % bromophenol blue, 0.1 % xylene cyanol, 2 mM EDTA) was added, and the netire sample electrophoresed in 15% denaturing polyacrylamide gel. Gels were fixed in 10% glycerol, 10% methanol, 10% glacial acetic acid with constant shaking followed by washing steps with 10% glycerol. The gels were then dried at 55 C for 3-5 hours.

The primers described in this experiment are described by Rychlik (Rychlik, (1995) *BioTechniques*, 18, p84-90). Primers may be synthesized or obtained as gel-filtration grade primers from Midland Certified Reagent Comapny (Midland Texas). The amplifications are either TAQ DNA polymerase-based (10 mM Tris-HCl pH 8.3, 1.5 mM  $MgCl_2$ , 50 mM KCl) or Pfu DNA polymerase-based based (20 mM Tris-HCl pH 8.3, 2.0 mM  $MgCl_2$ , 10 mM KCl, 10 mM  $(NH_4)_2SO_4$ , 0.1% Triton X-100, 0.1 mg/ml bovine serum albumin). The total nucleoside triphosphate (NTPs) concentration in the reactions is 0.8 mM, the primer concentration is 200 nM (unless otherwise stated) and the template amount is 0.25 ng of bacteriophage lambda DNA per 20  $\mu$ l reaction. Cycling parameters were 94°C preheating step for 5 minutes followed by 94°C denaturing step for 1 minute, 55°C annealing step for 2 minutes, and a 72°C extension step for 30 seconds to 1 minute and a final extension at 72°C for 10 minutes. Amplifications were generally 30 -45 in number.

Two regions in the bacteriophage lambda genome (GenBank Accession #J02459) were chosen as the priming sites for amplification. The 5'-primer has a stable GC-rich 3'-end: the 3' primer is chosen so that a 381 bp product will result. The 5' forward primer is H17: 5'-GAACGAAAACCCCCCGC. The 3'-reverse primer is RP17: 5'-GATCGCCCCCAAACACATA.

The amplified product was then tested for the presence of a polymorphism at position 31245. The following primer was used in four single

nucleotide extension assays; SNE17: 5'-GAACGAAAACCCCCCGC. The four single nucleotide extension assays were then carried as described above. All the reactions are then pooled and 5 µl of the pooled material was injected onto the HPLC column (SeraSep, San Jose, CA) without further purification.

5 HPLC was carried out using automated HPLC instrumentation (Rainin, Emeryville, CA, or, Hewlett Packard, Palo Alto, CA). Unpurified SNEA products which had been denatured for 3 minutes at 95 prior into injection into an HPLC were eluted with linear acetonitrile (ACN, J.T. Baker, NJ) gradient of 1.8%/minute at a flow rate of 0.9 ml/minute. The start and end points were adjusted according to the size of  
10 the SNEA product. The temperature required for the successful resolution of the SNEA molecules was 50°C. The effluent from the HPLC was then directed into a mass spectrometer (Hewlett Packard, Palo Alto, CA) for the detection of tags, with the results shown in TABLE 7.

15 TABLE 7

Tagged Primer	ddNTP Type	Retention Time	Extended?
SNE17-487	ddATP	2.5 minutes	no
SNE17-496	ddGTP	2.5 minutes	no
SNE17-503	ddCTP	4.6 minutes	yes
SNE17-555	ddTTP	2.5 minutes	no

The results of TABLE 7 indicate that the mass spectrometer tag (CMST) was detected at a retention time of 4.6 minutes indicating that the SNE17 primer was  
20 extended by one base (ddCTP) and therefore the polymorphism at position 31245 was in this case a "G". The SNE17-487, SNE17-496, and SNE17-555 tagged primers were not extended and their retention times on the HPLC was 2.5 minutes respectively.

25 EXAMPLE 27

PHORPHORAMIDITE CHEMISTRY FOR TAGGED MOLECULE SYNTHESIS

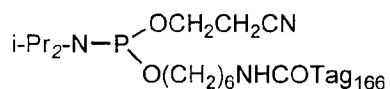
**Preparation of aminohexyl tailed tags**

As shown in Figure 26, preparation of pure aminohexyltailed tag (8) was  
30 achieved utilising the TFP methodology. The TFP ester was prepared by reaction of the

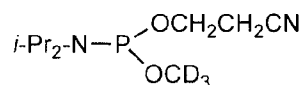
lithium salt of tag 166 (9) with excess TFP-TFA and Hunigs base in DMF. After workup a dichloromethane solution TFP ester was treated with 6-aminohexanol. A precipitate of (8) was formed immediately, and was isolated by filtration.

## 5 Preparation of tag phosphoramidites

Reaction of the aminohexyl tailed tag (5) with chlorophosphoramidite gave the desired phosphoramidite (17) with no phosphonate detected by mass spec analysis.



(17)



(18)

Excess phosphitylating agent was quenched with deuterated methanol (H3-methanol is usually used<sup>1</sup> but use of deuterated methanol in this case allows unambiguous identification of the origin of quench products) giving rise to (18). This material survived the subsequent aqueous workup and as it is itself a phosphoramidite it has the potential to interfere in subsequent reactions.

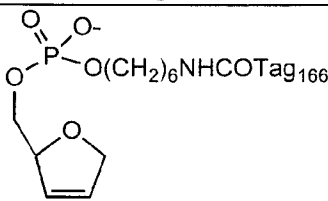
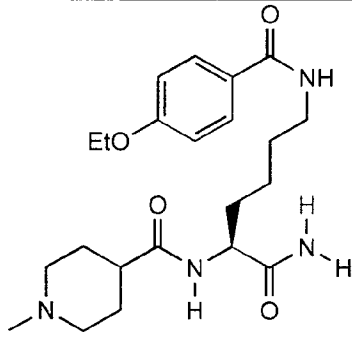
## Reaction of Phosphoramidite with Polymer Supported Thymidine

The thymidine is linked to controlled pore glass beads which are contained within a plastic cartridge. The cartridges used contain 1000 nM ( $1 \times 10^{-6}$  moles) of supported thymidine. The 5' hydroxyl of the base is protected as the dimethoxytrityl ether which is removed with 3% trichloroacetic acid in dichloromethane prior to reaction with the phosphoramidite (Figure 27). After reaction with the phosphoramidite the intermediate phosphite triester (19) is oxidised to the more stable phosphotriester (20). Concurrent removal of the cyanoethyl protecting group and cleavage from the solid support is carried out with ammonia to give the tagged thymidine (21).

Crude material from the phosphoramidite preparation was used in this sequence and the material cleaved from the support analysed by mass spectrometry. The major ions observed in positive mode along with assignments are shown in TABLE 8.

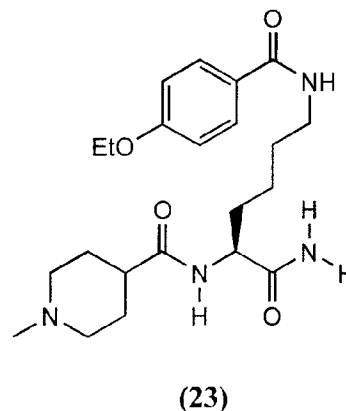
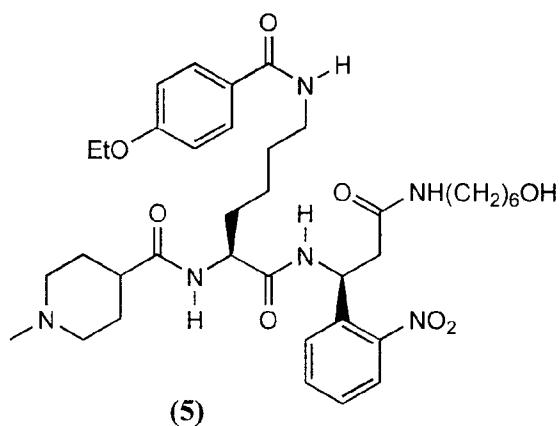
TABLE 8

212

Mass	Assignment
871	 Loss of thymine and water
711	$\text{O}(\text{CH}_2)_6\text{NHCOTag}_{166}$ Loss of thymidine and phosphate
419	 Photocleavage product.

It appears that under the conditions used the phosphate ester linkage is prone to fragmentation. Analysis in negative mode showed a peak at 338 amu corresponding to the compound **(22)** this arises from the reaction of the quench product **(18)** (Figure 28) which is a major contaminant of the crude.

The tag -phosphoramidite **(17)** could be separated from the non-polar contaminant **(18)** by silica gel chromatography (10% methanol / 89% dichloromethane / 1% triethylamine) and 115mg of **(17)** was obtained. This material was used in the reaction cycle to give **(21)** detected as its cleavage product **(5)** in the mass spec (HPLC analysis of the crude showed that **(5)** was not present thus the peak at 711 amu in the mass spec must have arisen by cleavage of the desired product **(21)** in the instrument). Iso present was unreacted thymidine indicating that the reaction had not gone to completion. The material was also passed through the UV cleavage flow system and the photocleavage product **(23)** detected by mass spectrometry.



Therefore, it is possible to prepare a mass spec tag phosphoramidite and  
 5 to purify and react it with a polymer bound substrate which can be subsequently  
 oxidised, deprotected and cleaved from the support to give a tagged base

1. "Oligonucleotide Synthesis, a Practical Approach" ed M.J.Gait, IRL Press 1990.
- 10 2. R.T Pon, *Tetrahedron lett*, 1991, 1715.
3. Available from Glen Research Cat no 10-1953-xx.
4. "Principles of Peptide Synthesis" M.Bodansky, Springer-Verlag 1984.
5. A.J Cocuzza, *Tetrahedron lett*, 1989, 6287.
- 15 6. N.N.Poluskin, B.Chen, L.W.Anderson and J.S Cohen, *J. Org. Chem*, 1993, **58**, 4606.

#### REFERENCES

- 20 Adams,S.P., Kavka,K.S., Wykes,E.J., Holder,S.B., and Galluppi,G.R. (1983). *J. Amer.Chem. Soc.* 105, 661-663.
- Andrus,A.,Efcavitch,J.W.,McBride,L.T., and Giusti,B. (1988). *Tetrahedron Lett.* 29, 861-864.
- Beaucage,S.L., and Caruthers,M.H. (1981). *Tetrahedron Lett.* 22,1859-1862.
- 25 Caruthers,M.H., Beaucage,S.L., Efcavitch,J.W., Fisher,E.F., Matteuci,M.D., and Stabinsky (1980). *Nucleic Acids Res.Symp.Ser.* 7, 215-223.

Ferrance, I.K., Eadie, J.S., and Ivarie, R.. (1989). *Nucleic Acids Res.* 17, 1232-1245.

Froehler, B.C., Ng, P.G., and Matteucci, M.D. (1986a). *Nucleic Acids Res.* 14, 5399-5407.

5 Froehler, B.C., and Matteucci, M.D. (1986b). *Tetrahedron Lett.* 27, 469-472.

Froehler, B.C. (1986c). *Tetrahedron Lett.* 27, 5575-5578.

Froehler, B.C., Ng, P.G., and Matteucci, M.D. (1988). *Nucleic Acids Res.* 16, 4831-4839.

10 Gaffney, B.L. and Jones, R.A.. (1988). *Tetrahedron Lett.* 29, 2619-2622.

Garegg, P.J., Regberg, T., Stawinski, J. and Stromberg, R. (1985). *Chem. Scripta* 25, 280-282.

Garegg, P.J., Regberg, T., Stawinski, J. and Stromberg, R. (1986a). *Chem. Scripta* 26, 59-62.

15 Garegg, P.J., Lidh, I., Regberg, T., Stawinski, J. and Stromberg, R. (1986b). *Tetrahedron Lett.* 27, 4051-4054.

Hall, R.H., Todd, A.R., and Webb, R.F.. (1957). *J. Chem. Soc.*, 3291-3296.

Letsinger, R.L., and Mahadevan, V.J. (1965). *J. Am. Chem. Soc.* 87, 3526-3227.

20 Letsinger, R.L., Finnan, J.L., Heavner, G.A. and Lunsford, W.B. (1975). *J. Amer. Chem. Soc.* 97, 3278.

Letsinger, R.L., Groody, E.P., and Tanaka, T. (1982). *J. Am. Chem. Soc.* 104, 6805-6806.

25 McBride, L.J., and Caruthers, M.H. (1983). *Tetrahedron Lett.* 24, 245-248.

Merrifield, R.B. (1963). *J. Am. Chem. Soc.* 85, 2149-2154.

Seliger, H. and Roesch, R. (1990). *DNA and Cell Biol.*, 9, 691-696.

Stein, C.A., Iversen, C.A., Subashinge, C., Cohen, J.S., Stec, W.J. and Zon, G. (1990). *Analytical Biochem.* 188, 11-16.

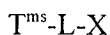
30

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention.

## Claims

WE CLAIM:

1. A compound of the formula:



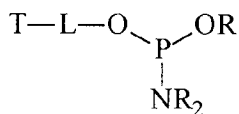
wherein,

$T^{ms}$  is an organic group detectable by mass spectrometry, comprising carbon, at least one of hydrogen and fluoride, and optional atoms selected from oxygen, nitrogen, sulfur, phosphorus and iodine;

L is an organic group which allows a unique  $T^{ms}$ -containing moiety to be cleaved from the remainder of the compound, wherein the  $T^{ms}$ -containing moiety comprises a functional group which supports a single ionized charge state when the compound is subjected to mass spectrometry and is tertiary amine, quaternary amine or organic acid;

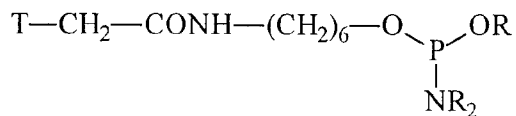
X is a functional group selected from phosphoramidite and H-phosphonate.

2. The compound of claim 1 wherein X is a phosphoramidite group such that T-L-X has the structure



wherein R is an alkyl group or a substituted alkyl group having one or more substituents selected from halogen and cyano, and the two R groups of  $NR_2$  may be bonded together to form a cycloalkyl group.

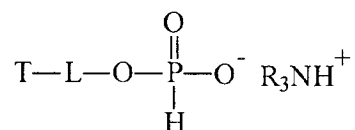
3. The compound of claim 2 wherein X is a phosphoramidite group such that T-L-X has the structure



and OR is  $OCH_2CH_2CN$  while  $NR_2$  is  $N(iso-propyl)_2$ .

4. The compound of claim 1 wherein X has an H-phosphonate group such that T-L-X has the structure





wherein R is a C<sub>1</sub>-C<sub>6</sub>alkyl group.

5. A compound according to claims 1-4 wherein T<sup>ms</sup> has a mass of from 15 to 10,000 daltons and a molecular formula of C<sub>1-500</sub>N<sub>0-100</sub>O<sub>0-100</sub>S<sub>0-10</sub>P<sub>0-10</sub>H<sub>α</sub>F<sub>β</sub>I<sub>δ</sub> wherein the sum of α, β and δ is sufficient to satisfy the otherwise unsatisfied valencies of the C, N and O atoms.

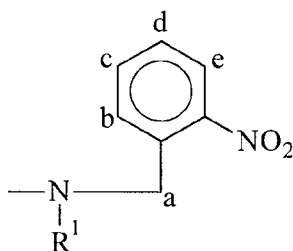
6. A compound according to claims 1-4 wherein T<sup>ms</sup> and L are bonded together through a functional group selected from amide, ester, ether, amine, sulfide, thioester, disulfide, thioether, urea, thiourea, carbamate, thiocarbamate, Schiff base, reduced Schiff base, imine, oxime, hydrazone, phosphate, phosphonate, phosphoramidate, phosphonamide, sulfonate, sulfonamide or carbon-carbon bond.

7. A compound according to claims 1-4 wherein L is selected from L<sup>hu</sup>, L<sup>acid</sup>, L<sup>base</sup>, L<sup>[O]</sup>, L<sup>[R]</sup>, L<sup>enz</sup>, L<sup>ele</sup>, L<sup>Δ</sup> and L<sup>ss</sup>, where actinic radiation, acid, base, oxidation, reduction, enzyme, electrochemical, thermal and thiol exchange, respectively, cause the T<sup>ms</sup>-containing moiety to be cleaved from the remainder of the molecule.

8. A compound according to claims 1-4 wherein L<sup>hu</sup> has the formula L<sup>1</sup>-L<sup>2</sup>-L<sup>3</sup>, wherein L<sup>2</sup> is a molecular fragment that absorbs actinic radiation to promote the cleavage of T<sup>ms</sup> from X, and L<sup>1</sup> and L<sup>3</sup> are independently a direct bond or an organic moiety, where L<sup>1</sup> separates L<sup>2</sup> from T<sup>ms</sup> and L<sup>3</sup> separates L<sup>2</sup> from X, and neither L<sup>1</sup> nor L<sup>3</sup> undergo bond cleavage when L<sup>2</sup> absorbs the actinic radiation.

9. A compound according to claim 8 wherein -L<sup>2</sup>-L<sup>3</sup> has the formula:

217

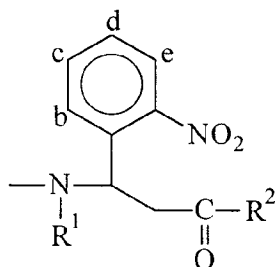


with one carbon atom at positions a, b, c, d or e being substituted with -  
 $L^3-X$  and optionally one or more of positions b, c, d or e being substituted with alkyl,  
 alkoxy, fluoride, chloride, hydroxyl, carboxylate or amide; and  $R^1$  is hydrogen or  
 5 hydrocarbyl.

10. A compound according to claim 8 wherein  $L^3$  is selected from a  
 direct bond, a hydrocarbylene, -O-hydrocarbylene, and hydrocarbylene-(O-  
 hydrocarbylene) $_n$ -H, and n is an integer ranging from 1 to 10.

10

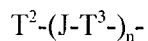
11. A compound according to claims 1-4 wherein -L-X has the  
 formula:



wherein one or more of positions b, c, d or e is substituted with  
 15 hydrogen, alkyl, alkoxy, fluoride, chloride, hydroxyl, carboxylate or amide;  $R^1$  is  
 hydrogen or hydrocarbyl, and  $R^2$  terminates in an "X" group.

12. A compound according to claims 1-4 wherein  $T^{ms}$  has the  
 formula:

20



$T^2$  is an organic moiety formed from carbon and one or more of  
 hydrogen, fluoride, iodide, oxygen, nitrogen, sulfur and phosphorous, having a mass of  
 15 to 500 daltons;

T<sup>3</sup> is an organic moiety formed from carbon and one or more of hydrogen, fluoride, iodide, oxygen, nitrogen, sulfur and phosphorous, having a mass of 50 to 1000 daltons;

J is a direct bond or a functional group selected from amide, ester, amine, sulfide, ether, thioester, disulfide, thioether, urea, thiourea, carbamate, thiocarbamate, Schiff base, reduced Schiff base, imine, oxime, hydrazone, phosphate, phosphonate, phosphoramidate, phosphonamide, sulfonate, sulfonamide or carbon-carbon bond; and

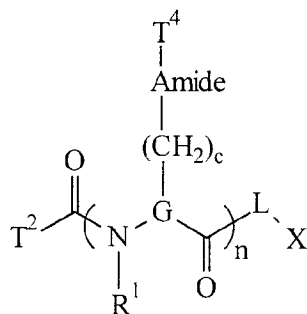
n is an integer ranging from 1 to 50, and when n is greater than 1, each T<sup>3</sup> and J is independently selected.

13. A compound according to claim 12 wherein T<sup>2</sup> is selected from hydrocarbyl, hydrocarbyl-O-hydrocarbylene, hydrocarbyl-S-hydrocarbylene, hydrocarbyl-NH-hydrocarbylene, hydrocarbyl-amide-hydrocarbylene, N-(hydrocarbyl)hydrocarbylene, N,N-di(hydrocarbyl)hydrocarbylene, hydrocarbylacyl-hydrocarbylene, heterocyclylhydrocarbyl wherein the heteroatom(s) are selected from oxygen, nitrogen, sulfur and phosphorous, substituted heterocyclylhydrocarbyl wherein the heteroatom(s) are selected from oxygen, nitrogen, sulfur and phosphorous and the substituents are selected from hydrocarbyl, hydrocarbyl-O-hydrocarbylene, hydrocarbyl-NH-hydrocarbylene, hydrocarbyl-S-hydrocarbylene, N-(hydrocarbyl)hydrocarbylene, N,N-di(hydrocarbyl)hydrocarbylene and hydrocarbylacyl-hydrocarbylene, as well as derivatives of any of the foregoing wherein one or more hydrogens is replaced with an equal number of fluorides.

14. A compound according to claim 12 wherein T<sup>3</sup> has the formula -G(R<sup>2</sup>)-, G is C<sub>1-6</sub> alkylene having a single R<sup>2</sup> substituent, and R<sup>2</sup> is selected from alkyl, alkenyl, alkynyl, cycloalkyl, aryl-fused cycloalkyl, cycloalkenyl, aryl, aralkyl, aryl-substituted alkenyl or alkynyl, cycloalkyl-substituted alkyl, cycloalkenyl-substituted cycloalkyl, biaryl, alkoxy, alkenoxy, alkynoxy, aralkoxy, aryl-substituted alkenoxy or alkynoxy, alkylamino, alkenylamino or alkynylamino, aryl-substituted alkylamino, aryl-substituted alkenylamino or alkynylamino, aryloxy, arylamino, N-alkylurea-substituted alkyl, N-arylurea-substituted alkyl, alkylcarbonylamino-substituted alkyl, aminocarbonyl-substituted alkyl, heterocyclyl, heterocyclyl-substituted alkyl, heterocyclyl-substituted amino, carboxyalkyl substituted aralkyl, oxocarbocyclyl-fused aryl and heterocyclylalkyl; cycloalkenyl, aryl-substituted alkyl and, aralkyl, hydroxy-substituted alkyl, alkoxy-substituted alkyl, aralkoxy-substituted alkyl, alkoxy-substituted alkyl, aralkoxy-substituted alkyl, amino-

- substituted alkyl, (aryl-substituted alkyloxycarbonylamino)-substituted alkyl, thiol-substituted alkyl, alkylsulfonyl-substituted alkyl, (hydroxy-substituted alkylthio)-substituted alkyl, thioalkoxy-substituted alkyl, hydrocarbylacylamino-substituted alkyl, heterocyclylacylamino-substituted alkyl, hydrocarbyl-substituted-
- 5 heterocyclylacylamino-substituted alkyl, alkylsulfonylamino-substituted alkyl, arylsulfonylamino-substituted alkyl, morpholino-alkyl, thiomorpholino-alkyl, morpholino carbonyl-substituted alkyl, thiomorpholinocarbonyl-substituted alkyl, [N-(alkyl, alkenyl or alkynyl)- or N,N-[dialkyl, dialkenyl, dialkynyl or (alkyl, alkenyl)-amino]carbonyl-substituted alkyl, heterocyclylaminocarbonyl,
- 10 heterocyclylalkyleneaminocarbonyl, heterocyclylaminocarbonyl-substituted alkyl, heterocyclylalkyleneaminocarbonyl-substituted alkyl, N,N-[dialkyl]alkyleneaminocarbonyl, N,N-[dialkyl]alkyleneaminocarbonyl-substituted alkyl, alkyl-substituted heterocyclylcarbonyl, alkyl-substituted heterocyclylcarbonyl-alkyl, carboxyl-substituted alkyl, dialkylamino-substituted acylaminoalkyl and amino
- 15 acid side chains selected from arginine, asparagine, glutamine, S-methyl cysteine, methionine and corresponding sulfoxide and sulfone derivatives thereof, glycine, leucine, isoleucine, allo-isoleucine, tert-leucine, norleucine, phenylalanine, tyrosine, tryptophan, proline, alanine, ornithine, histidine, glutamine, valine, threonine, serine, aspartic acid, beta-cyanoalanine, and allothreonine; alynyl and heterocyclylcarbonyl,
- 20 aminocarbonyl, amido, mono- or dialkylaminocarbonyl, mono- or diarylaminoacarbonyl, alkylarylaminocarbonyl, diarylaminoacarbonyl, mono- or diacylaminoacarbonyl, aromatic or aliphatic acyl, alkyl optionally substituted by substituents selected from amino, carboxy, hydroxy, mercapto, mono- or dialkylamino, mono- or diarylamino, alkylarylamino, diarylamino, mono- or diacylamino, alkoxy, alkenoxy, aryloxy,
- 25 thioalkoxy, thioalkenoxo, thioalkynoxo, thioaryloxy and heterocyclyl.

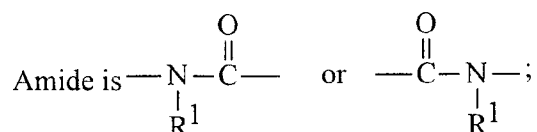
15. A compound according to claim 12 having the formula:



wherein

G is  $(\text{CH}_2)_{1-6}$  wherein a hydrogen on one and only one of the  $\text{CH}_2$  groups is replaced with  $-(\text{CH}_2)_c\text{-Amide-T}^4$ ;

$\text{T}^2$  and  $\text{T}^4$  are organic moieties of the formula  $\text{C}_{1-25}\text{N}_{0-9}\text{O}_{0-9}\text{H}_\alpha\text{F}_\beta$  wherein the sum of  $\alpha$  and  $\beta$  is sufficient to satisfy the otherwise unsatisfied valencies of the C,  
 5 N, and O atoms;



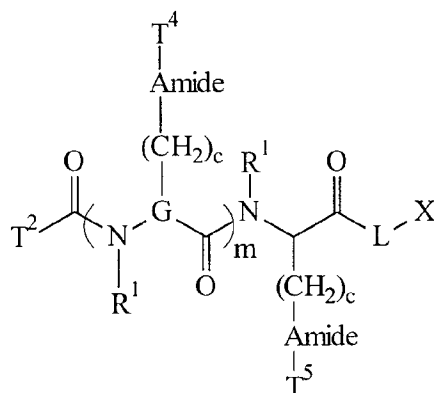
$\text{R}^1$  is hydrogen or  $\text{C}_{1-10}$  alkyl;

c is an integer ranging from 0 to 4;

X is defined according to claim 1; and

10 n is an integer ranging from 1 to 50 such that when n is greater than 1,  
 G, c, Amide,  $\text{R}^1$  and  $\text{T}^4$  are independently selected.

16. A compound according to claim 15 having the formula:

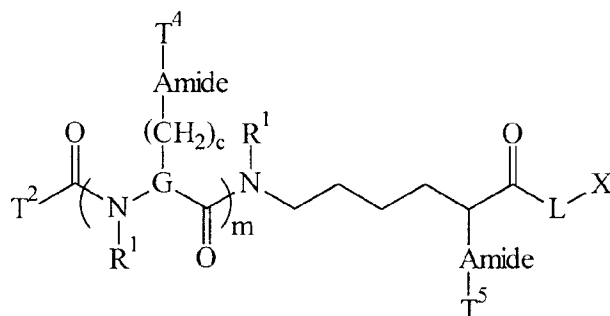


15

wherein  $\text{T}^5$  is an organic moiety of the formula  $\text{C}_{1-25}\text{N}_{0-9}\text{O}_{0-9}\text{H}_\alpha\text{F}_\beta$  wherein the sum of  $\alpha$  and  $\beta$  is sufficient to satisfy the otherwise unsatisfied valencies of the C, N, and O atoms; and  $\text{T}^5$  includes a tertiary or quaternary amine or an organic acid; and  
 20 m is an integer ranging from 0-49.

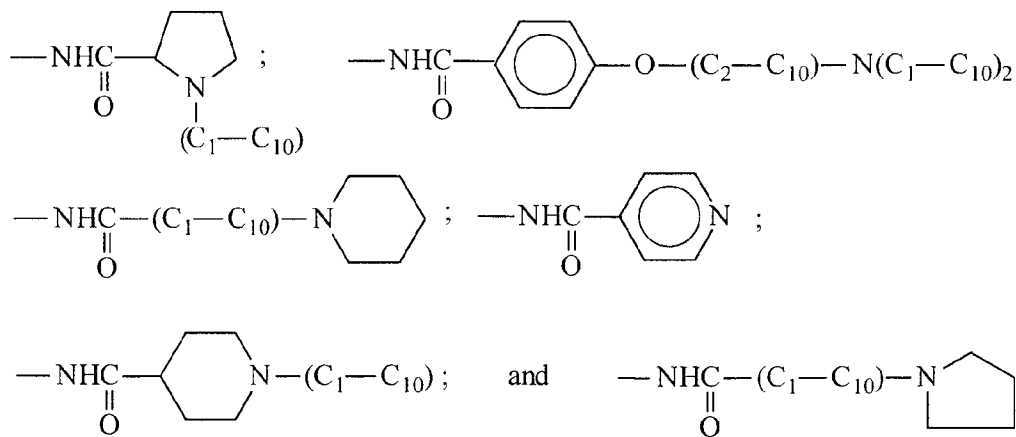
20

17. A compound according to claim 15 having the formula:

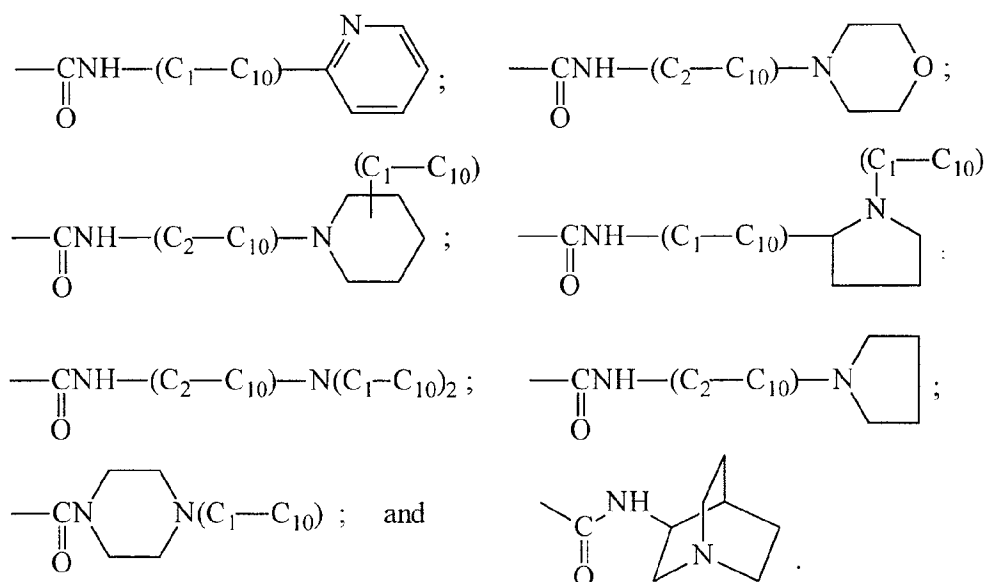


wherein T<sup>5</sup> is an organic moiety of the formula C<sub>1-25</sub>N<sub>0-9</sub>O<sub>0-9</sub>H<sub>α</sub>F<sub>β</sub> wherein the sum of α and β is sufficient to satisfy the otherwise unsatisfied valencies of the C, N, and O atoms; and T<sup>5</sup> includes a tertiary or quaternary amine or an organic acid; and m is an integer ranging from 0-49.

18. A compound according to any one of claims 16 and 17 wherein -Amide-T<sup>5</sup> is selected from:



19. A compound according to any of claims 16 and 17 wherein -Amide-T<sup>5</sup> is selected from:



20. A compound according to any one of claims 14-16 wherein T<sup>2</sup> has the structure which results when one of the following organic acids is condensed with an amine group to form T<sup>2</sup>-C(=O)-N(R<sup>1</sup>):
- Formic acid, Acetic acid, Propionic acid, Propionic acid, Fluoroacetic acid, 2-Butynoic acid, Cyclopropanecarboxylic acid, Butyric acid, Methoxyacetic acid, Difluoroacetic acid, 4-Pentynoic acid, Cyclobutanecarboxylic acid, 3,3-Dimethylacrylic acid, Valeric acid, N,N-Dimethylglycine, N-Formyl-Gly-OH, Ethoxyacetic acid, (Methylthio)acetic acid, Pyrrole-2-carboxylic acid, 3-Furoic acid, Isoxazole-5-carboxylic acid, trans-3-Hexenoic acid, Trifluoroacetic acid, Hexanoic acid, Ac-Gly-OH, 2-Hydroxy-2-methylbutyric acid, Benzoic acid, Nicotinic acid, 2-Pyrazinecarboxylic acid, 1-Methyl-2-pyrrolicarboxylic acid, 2-Cyclopentene-1-acetic acid, Cyclopentylacetic acid, (S)-(-)-2-Pyrrolidone-5-carboxylic acid, N-Methyl-L-proline, Heptanoic acid, Ac-b-Ala-OH, 2-Ethyl-2-hydroxybutyric acid, 2-(2-Methoxyethoxy)acetic acid, p-Toluic acid, 6-Methylnicotinic acid, 5-Methyl-2-pyrazinecarboxylic acid, 2,5-Dimethylpyrrole-3-carboxylic acid, 4-Fluorobenzoic acid, 3,5-Dimethylisoxazole-4-carboxylic acid, 3-Cyclopentylpropionic acid, Octanoic acid, N,N-Dimethylsuccinamic acid, Phenylpropionic acid, Cinnamic acid, 4-Ethylbenzoic acid, p-Anisic acid, 1,2,5-Trimethylpyrrole-3-carboxylic acid, 3-Fluoro-4-methylbenzoic acid, Ac-DL-Propargylglycine, 3-(Trifluoromethyl)butyric acid, 1-Piperidinepropionic acid, N-Acetylproline, 3,5-Difluorobenzoic acid, Ac-L-Val-OH, Indole-2-carboxylic acid, 2-Benzofurancarboxylic acid, Benzotriazole-5-carboxylic acid, 4-n-Propylbenzoic acid, 3-Dimethylaminobenzoic acid, 4-Ethoxybenzoic acid, 4-(Methylthio)benzoic acid, N-(2-Furoyl)glycine, 2-(Methylthio)nicotinic acid, 3-Fluoro-4-methoxybenzoic acid, Tfa-

- Gly-OH, 2-Napthoic acid, Quinaldic acid, Ac-L-Ile-OH, 3-Methylindene-2-carboxylic acid, 2-Quinoxalinecarboxylic acid, 1-Methylindole-2-carboxylic acid, 2,3,6-Trifluorobenzoic acid, N-Formyl-L-Met-OH, 2-[2-(2-Methoxyethoxy)ethoxy]acetic acid, 4-n-Butylbenzoic acid, N-Benzoylglycine, 5-Fluoroindole-2-carboxylic acid, 4-n-Propoxybenzoic acid, 4-Acetyl-3,5-dimethyl-2-pyrrolecarboxylic acid, 3,5-Dimethoxybenzoic acid, 2,6-Dimethoxynicotinic acid, Cyclohexanepentanoic acid, 2-Naphthylacetic acid, 4-(1H-Pyrrol-1-yl)benzoic acid, Indole-3-propionic acid, m-Trifluoromethylbenzoic acid, 5-Methoxyindole-2-carboxylic acid, 4-Pentylbenzoic acid, Bz-b-Ala-OH, 4-Diethylaminobenzoic acid, 4-n-Butoxybenzoic acid, 3-Methyl-5-CF<sub>3</sub>-isoxazole-4-carboxylic acid, (3,4-Dimethoxyphenyl)acetic acid, 4-Biphenylcarboxylic acid, Pivaloyl-Pro-OH, Octanoyl-Gly-OH, (2-Naphthoxy)acetic acid, Indole-3-butyric acid, 4-(Trifluoromethyl)phenylacetic acid, 5-Methoxyindole-3-acetic acid, 4-(Trifluoromethoxy)benzoic acid, Ac-L-Phe-OH, 4-Pentyloxybenzoic acid, Z-Gly-OH, 4-Carboxy-N-(fur-2-ylmethyl)pyrrolidin-2-one, 3,4-Diethoxybenzoic acid, 2,4-Dimethyl-5-CO<sub>2</sub>Et-pyrrole-3-carboxylic acid, N-(2-Fluorophenyl)succinamic acid, 3,4,5-Trimethoxybenzoic acid, N-Phenylanthranilic acid, 3-Phenoxybenzoic acid, Nonanoyl-Gly-OH, 2-Phenoxypyridine-3-carboxylic acid, 2,5-Dimethyl-1-phenylpyrrole-3-carboxylic acid, trans-4-(Trifluoromethyl)cinnamic acid, (5-Methyl-2-phenyloxazol-4-yl)acetic acid, 4-(2-Cyclohexenyloxy)benzoic acid, 5-Methoxy-2-methylindole-3-acetic acid, trans-4-Cotininecarboxylic acid, Bz-5-Aminovaleric acid, 4-Hexyloxybenzoic acid, N-(3-Methoxyphenyl)succinamic acid, Z-Sar-OH, 4-(3,4-Dimethoxyphenyl)butyric acid, Ac-o-Fluoro-DL-Phe-OH, N-(4-Fluorophenyl)glutaramic acid, 4'-Ethyl-4-biphenylcarboxylic acid, 1,2,3,4-Tetrahydroacridinecarboxylic acid, 3-Phenoxyphenylacetic acid, N-(2,4-Difluorophenyl)succinamic acid, N-Decanoyl-Gly-OH, (+)-6-Methoxy-a-methyl-2-naphthaleneacetic acid, 3-(Trifluoromethoxy)cinnamic acid, N-Formyl-DL-Trp-OH, (R)-(+)-a-Methoxy-a-(trifluoromethyl)phenylacetic acid, Bz-DL-Leu-OH, 4-(Trifluoromethoxy)phenoxyacetic acid, 4-Heptyloxybenzoic acid, 2,3,4-Trimethoxycinnamic acid, 2,6-Dimethoxybenzoyl-Gly-OH, 3-(3,4,5-Trimethoxyphenyl)propionic acid, 2,3,4,5,6-Pentafluorophenoxyacetic acid, N-(2,4-Difluorophenyl)glutaramic acid, N-Undecanoyl-Gly-OH, 2-(4-Fluorobenzoyl)benzoic acid, 5-Trifluoromethoxyindole-2-carboxylic acid, N-(2,4-Difluorophenyl)diglycolamic acid, Ac-L-Trp-OH, Tfa-L-Phenylglycine-OH, 3-Iodobenzoic acid, 3-(4-n-Pentylbenzoyl)propionic acid, 2-Phenyl-4-quinolinecarboxylic acid, 4-Octyloxybenzoic acid, Bz-L-Met-OH, 3,4,5-Triethoxybenzoic acid, N-Lauroyl-Gly-OH, 3,5-Bis(trifluoromethyl)benzoic acid, Ac-5-Methyl-DL-Trp-OH, 2-Iodophenylacetic acid, 3-Iodo-4-methylbenzoic acid, 3-(4-n-Hexylbenzoyl)propionic acid, N-Hexanoyl-L-Phe-



OH, 4-Nonyloxybenzoic acid, 4'-(Trifluoromethyl)-2-biphenylcarboxylic acid, Bz-L-Phe-OH, N-Tridecanoyl-Gly-OH, 3,5-Bis(trifluoromethyl)phenylacetic acid, 3-(4-n-Heptylbenzoyl)propionic acid, N-Hepytanoyl-L-Phe-OH, 4-Decyloxybenzoic acid, N-( $\alpha,\alpha,\alpha$ -trifluoro-m-tolyl)anthranilic acid, Niflumic acid, 4-(2-Hydroxyhexafluoroisopropyl)benzoic acid, N-Myristoyl-Gly-OH, 3-(4-n-Octylbenzoyl)propionic acid, N-Octanoyl-L-Phe-OH, 4-Undecyloxybenzoic acid, 3-(3,4,5-Trimethoxyphenyl)propionyl-Gly-OH, 8-Iodonaphthoic acid, N-Pentadecanoyl-Gly-OH, 4-Dodecyloxybenzoic acid, N-Palmitoyl-Gly-OH, and N-Stearoyl-Gly-OH.

21. A method for determining the presence of a single nucleotide polymorphism in a nucleic acid target comprising:
- a) amplifying a sequence of a nucleic acid target containing a single nucleotide polymorphism;
  - b) generating a single strand form of the target;
  - c) combining a tagged nucleic acid probe with the amplified target nucleic acid molecules under conditions and for a time sufficient to permit hybridization of said tagged nucleic acid probe to complementary amplified selected target nucleic acid molecules, wherein said tag is correlative with a particular single nucleotide polymorphism and is detectable by spectrometry or potentiometry;
  - d) separating unhybridized tagged probe from hybridized tagged probe by a sizing methodology;
  - e) cleaving said tag from said probe; and
  - f) detecting said tag by spectrometry or potentiometry, and determining the presence of said single nucleotide polymorphism.

22. A method for determining the presence of a single nucleotide polymorphism in a nucleic acid target comprising:
- a) amplifying a sequence of a nucleic acid target containing a single nucleotide polymorphism;
  - b) combining a tagged nucleic acid primer with the amplified target nucleic acid molecules under conditions and for a time sufficient to permit annealing of said tagged nucleic acid primer to complementary amplified selected target nucleic acid molecules, wherein the oligonucleotide primer has a 3'-most base complementary to the wildtype sequence or the single nucleotide polymorphism, wherein said tag is correlative with a particular single nucleotide polymorphism and is detectable by spectrometry or potentiometry;

c) extending the primer wherein a complementary strand to the target is synthesized when the 3'-most base of the primer is complementary to the target;

d) separating unextended tagged primer from extended tagged primer by a sizing methodology;

5 e) cleaving said tag from said primers or extended primers; and

f) detecting said tag by spectrometry or potentiometry, and determining therefrom the presence of said single nucleotide polymorphism.

23. A method for determining the quantity of a specific mRNA molecule in a nucleic acid population comprising:

10 a) converting an RNA population into a cDNA population;

b) adding a single strand nucleic acid (internal standard) containing a plurality of single nucleotide polymorphisms, that is otherwise identical to said cDNA target;

c) amplifying a specific sequence of said cDNA target;

15 d) coamplifying the internal standard, wherein said internal standard is the same length as the cDNA amplicon;

e) generating a single strand form of the target;

20 f) combining a set of tagged nucleic acid probes with the amplified target cDNA and amplified internal standard under conditions and for a time sufficient to permit hybridization of said tagged nucleic acid probe to complementary selected target cDNA and internal standard sequences, wherein said tag is correlative with a particular cDNA sequence and a second tag is correlative with the internal standard, and is detectable by spectrometry or potentiometry;

25 g) separating unhybridized tagged probe from hybridized tagged probe by a sizing methodology;

h) cleaving said tag from said probes;

i) detecting said tags by spectrometry or potentiometry; and

30 j) taking the ratio of tag correlated to cDNA to tag correlated with the internal standard, and determining therefrom the quantity of said cDNA, thereby determining the quantity of the specific mRNA in a nucleic acid population.

24. A method for determining the quantity of a single nucleotide polymorphism in a nucleic acid target comprising:

a) amplifying a sequence of a nucleic acid target containing a single nucleotide polymorphism;

b) generating a single strand form of the target;

c) combining a tagged nucleic acid probe with the amplified target nucleic acid molecules under conditions and for a time sufficient to permit hybridization of said tagged nucleic acid probe to complementary amplified selected target nucleic acid molecules, wherein said tag is correlative with a particular single nucleotide polymorphism and is detectable by spectrometry or potentiometry;

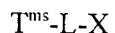
d) separating unhybridized tagged probe from hybridized tagged probe by a sizing methodology;

e) cleaving said tag from said probes;

f) detecting said tags by spectrometry or potentiometry; and

j) taking the ratio of tag correlated to the wild type polymorphism to the tag correlated with the mutant polymorphism, and determining therefrom the quantity of said polymorphism.

25. The method of any of claims 21-24 wherein the tagged nucleic acid has the structure



wherein,

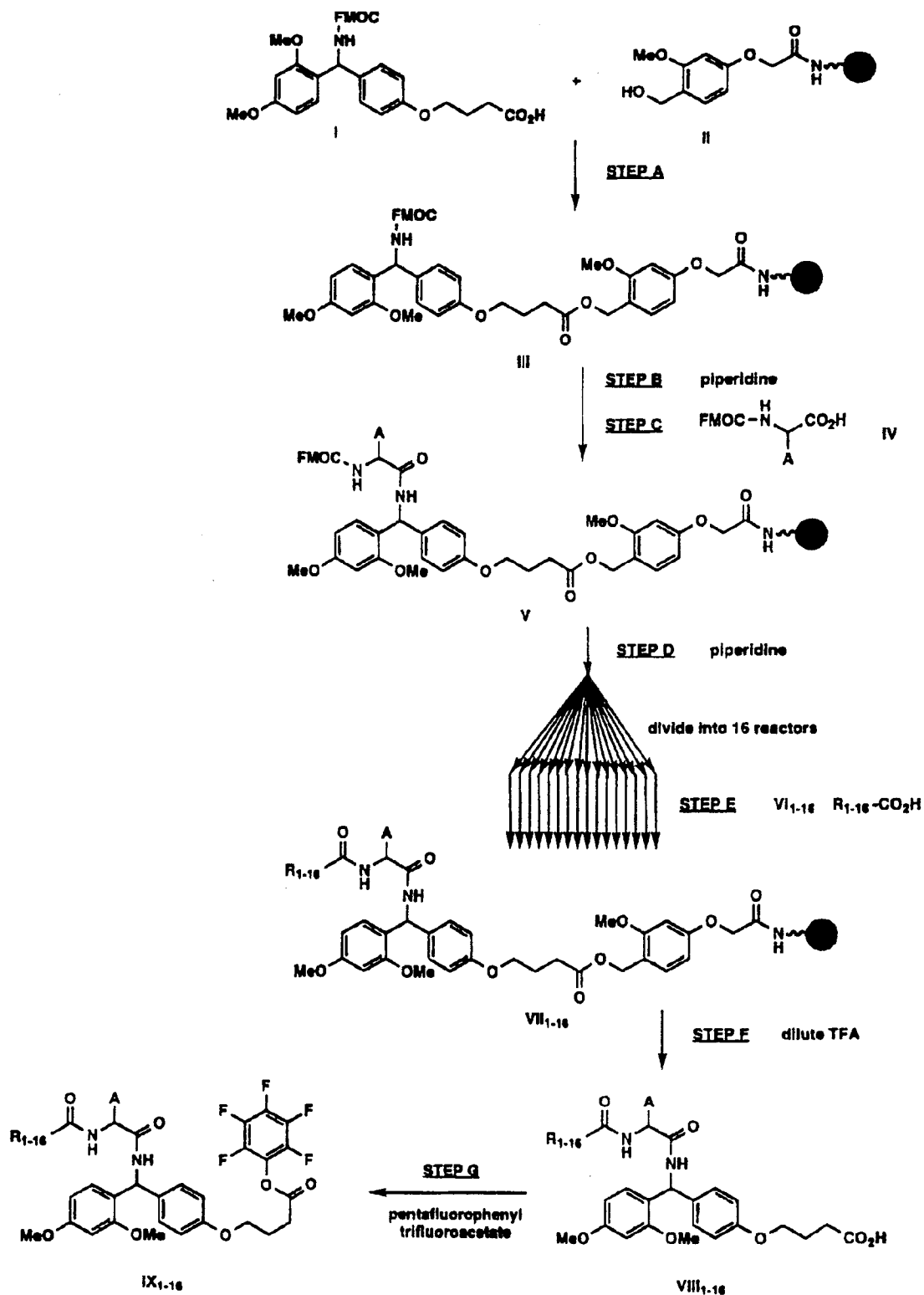
$T^{\text{ms}}$  is an organic group detectable by mass spectrometry, comprising carbon, at least one of hydrogen and fluoride, and optional atoms selected from oxygen, nitrogen, sulfur, phosphorus and iodine;

L is an organic group which allows a unique  $T^{\text{ms}}$ -containing moiety to be cleaved from the remainder of the compound, wherein the  $T^{\text{ms}}$ -containing moiety comprises a functional group which supports a single ionized charge state when the compound is subjected to mass spectrometry and is tertiary amine, quaternary amine or organic acid; and

X is nucleic acid.

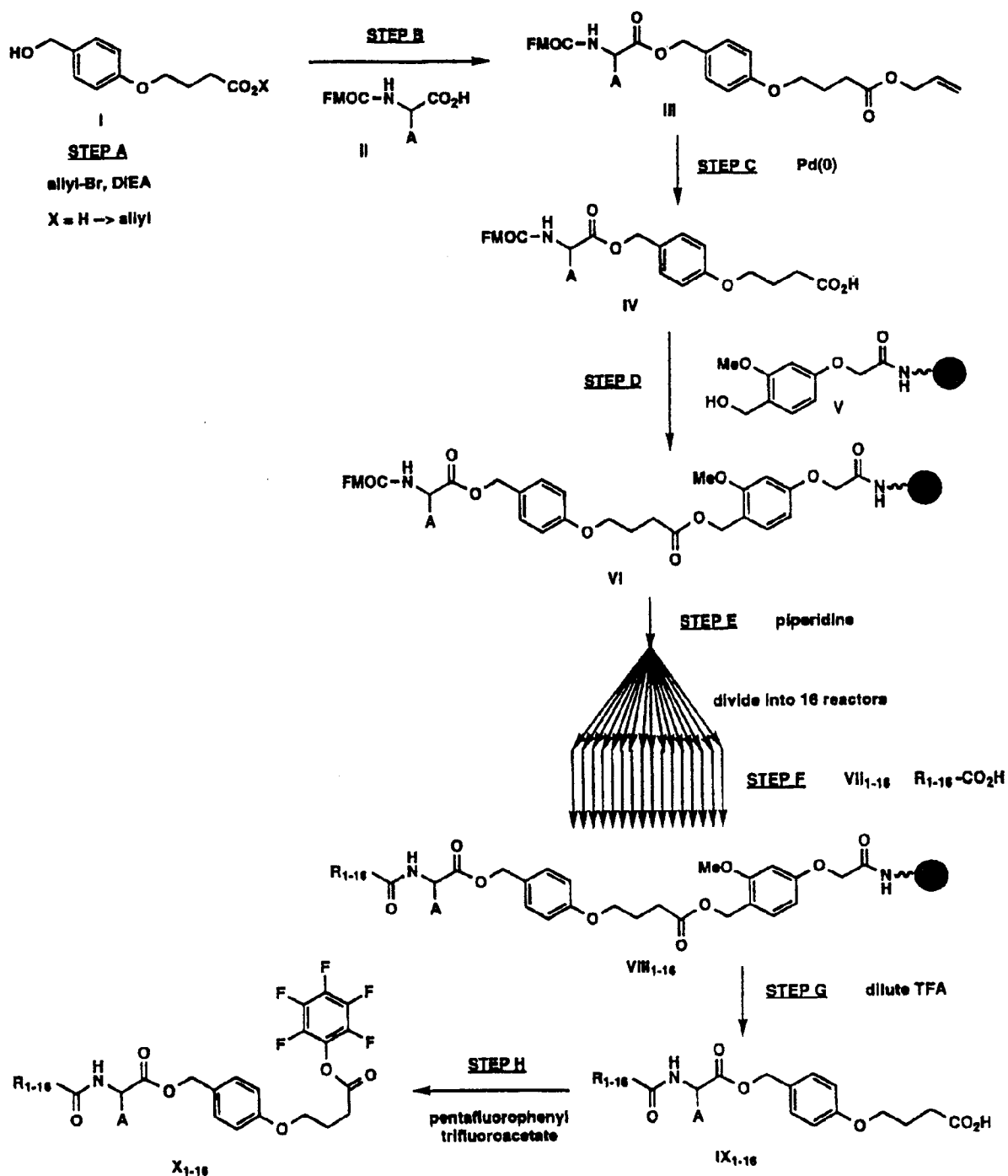
1/36

## FIGURE 1



2/36

## FIGURE 2



3/36

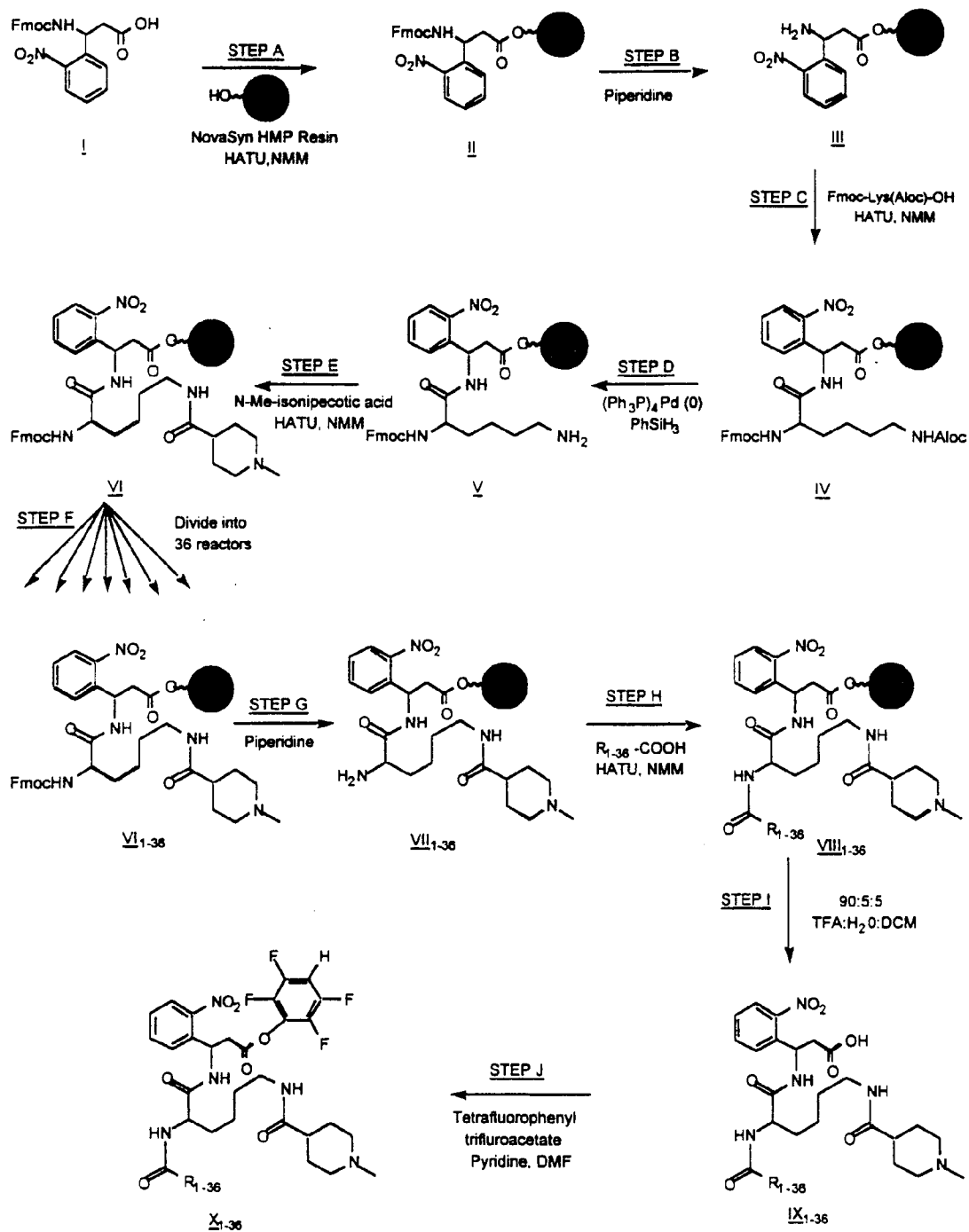


FIGURE 3

4/36

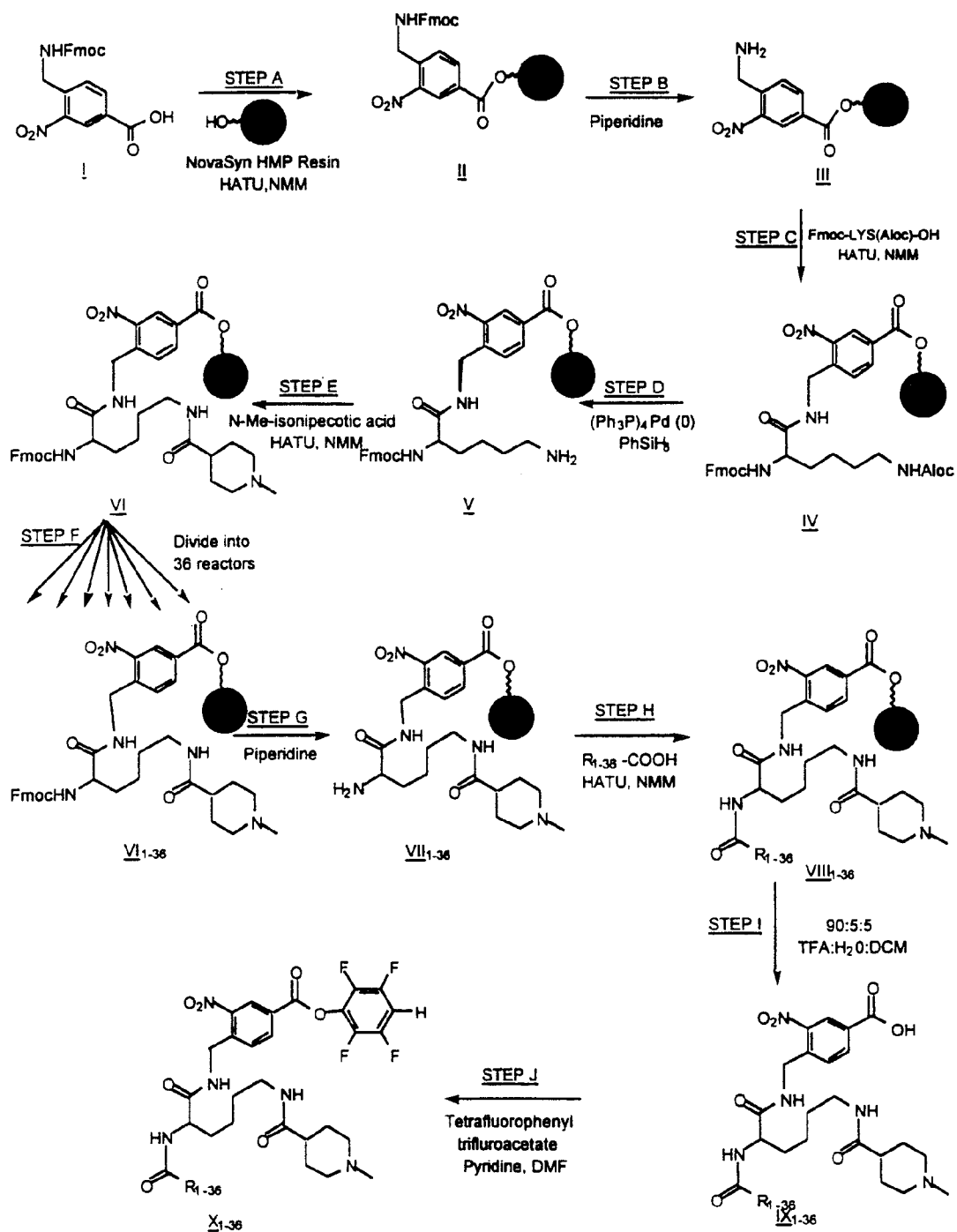


FIGURE 4

5/36

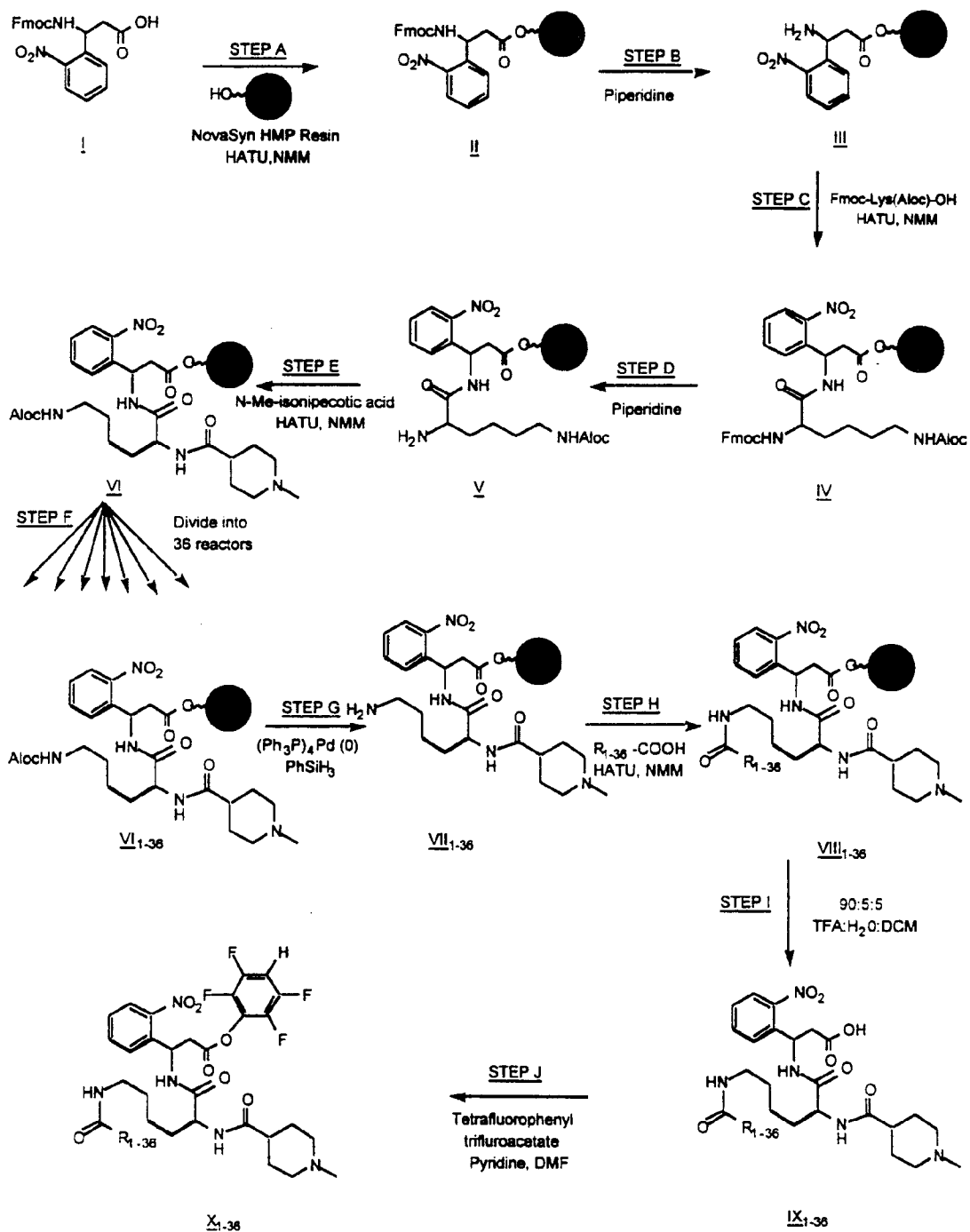


FIGURE 5



6/36

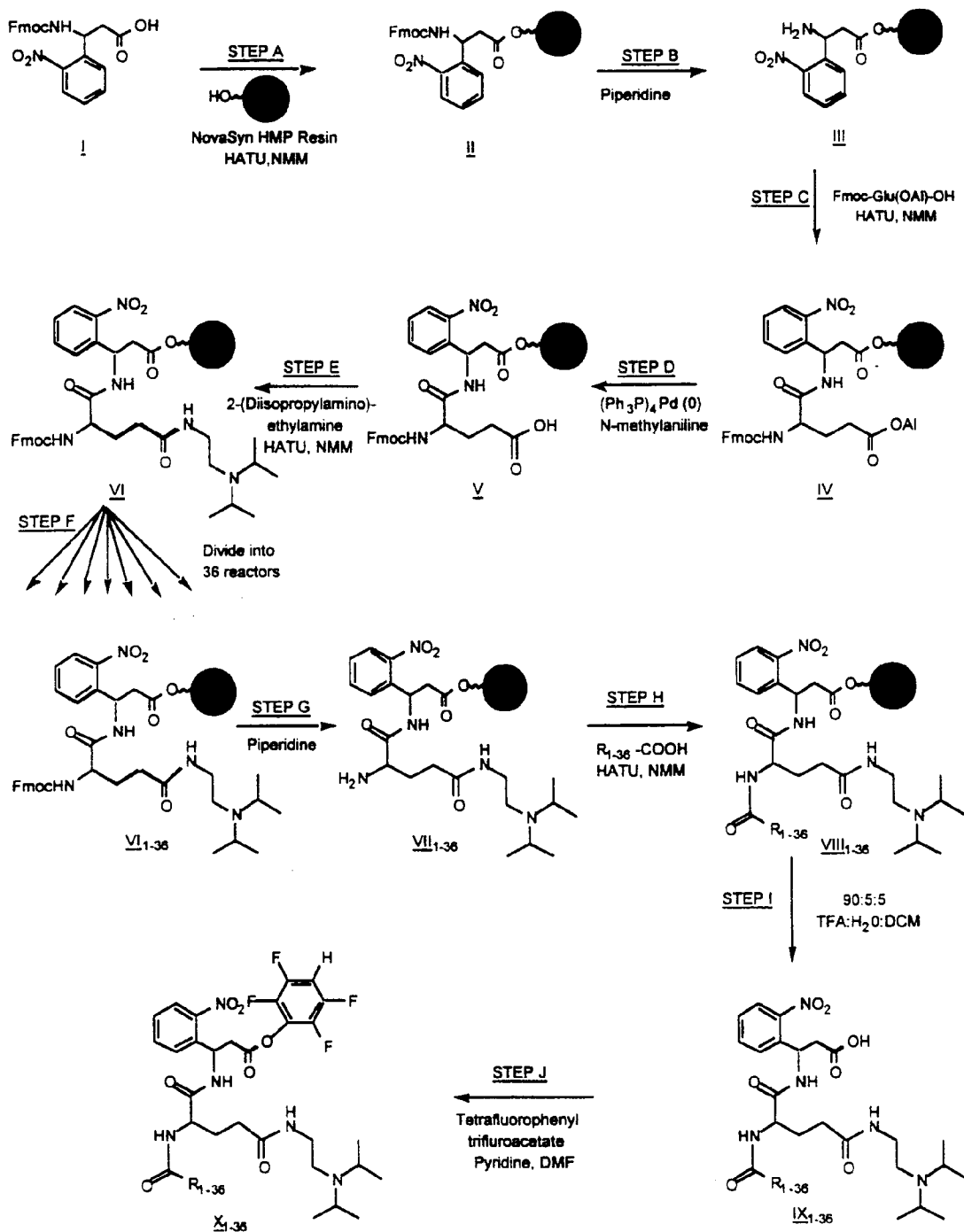


FIGURE 6

7/36

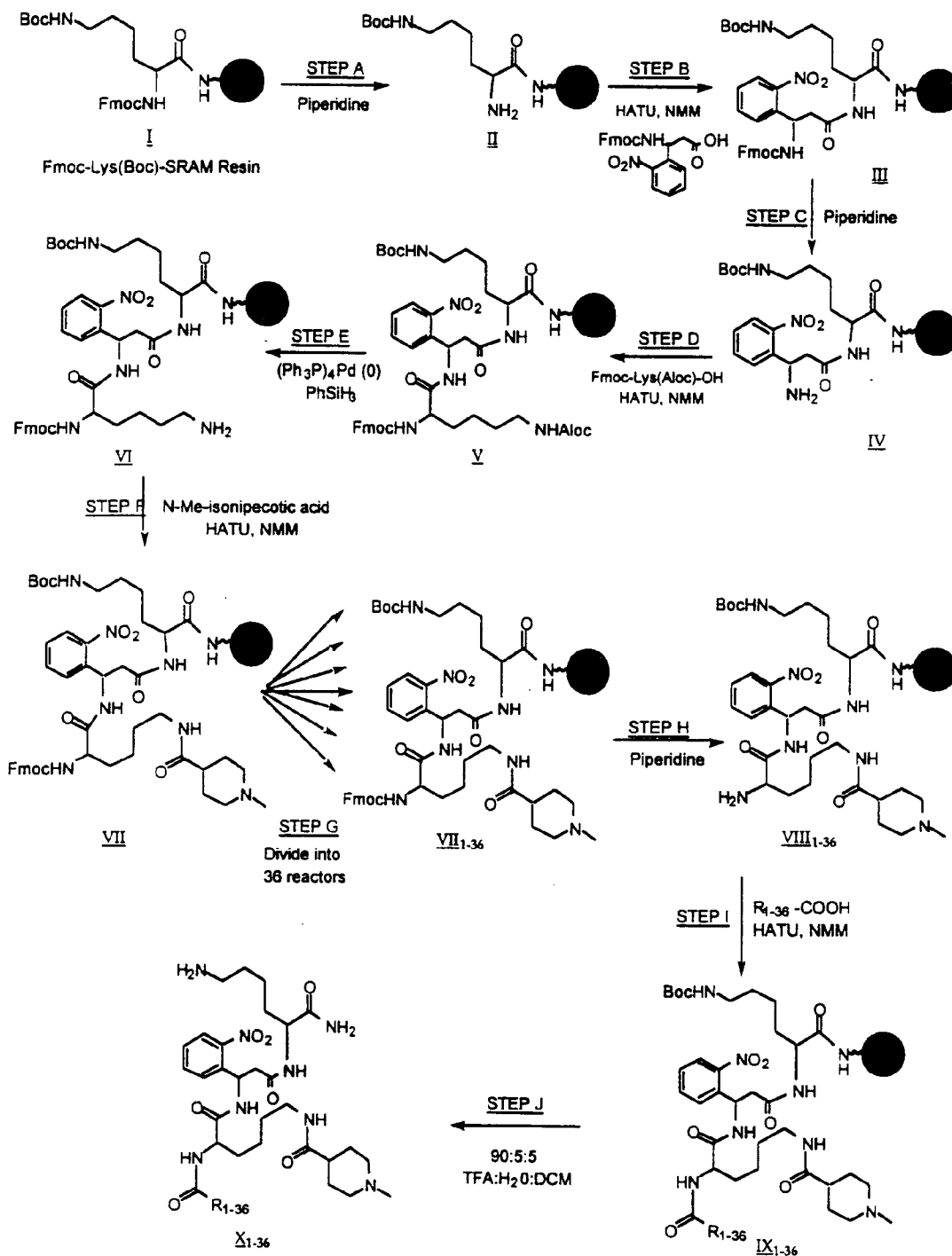
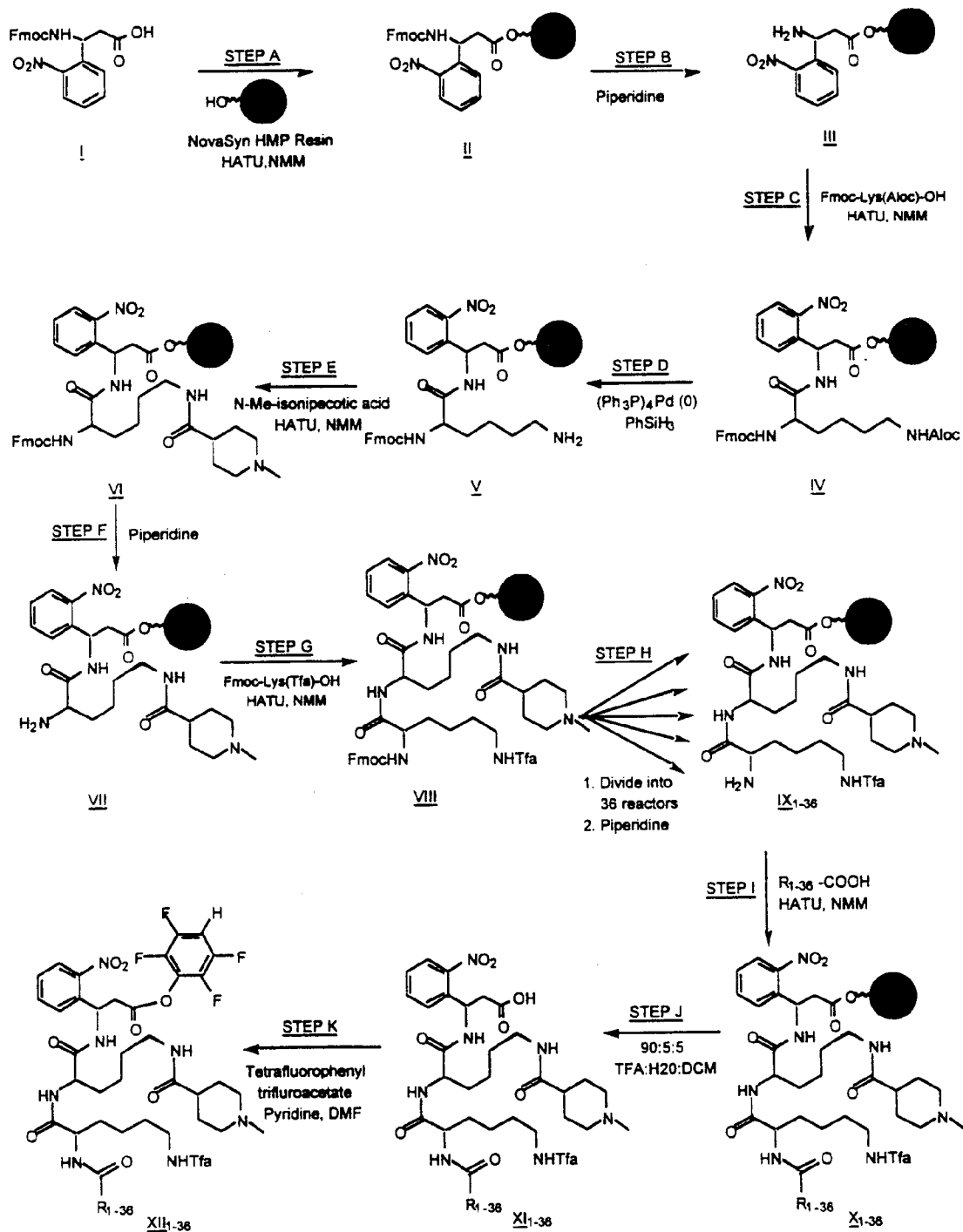


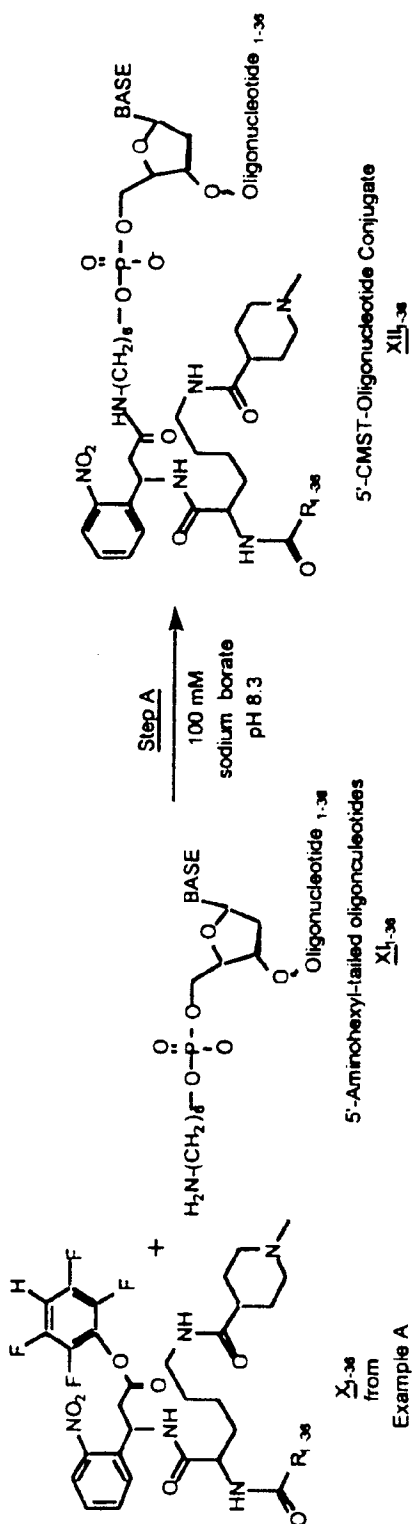
FIGURE 7

8/36



## FIGURE 8

9/36



## FIGURE 9

10/36

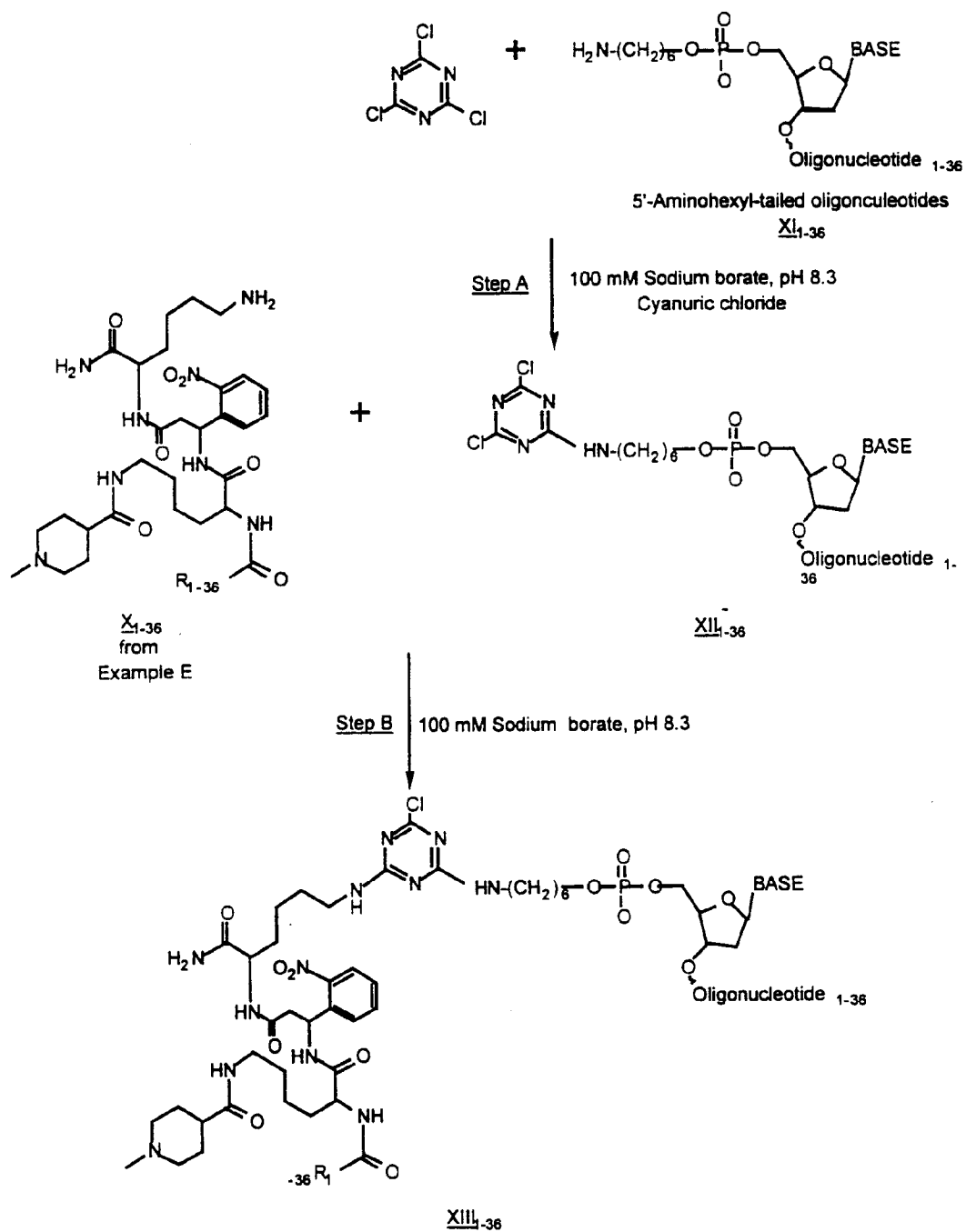
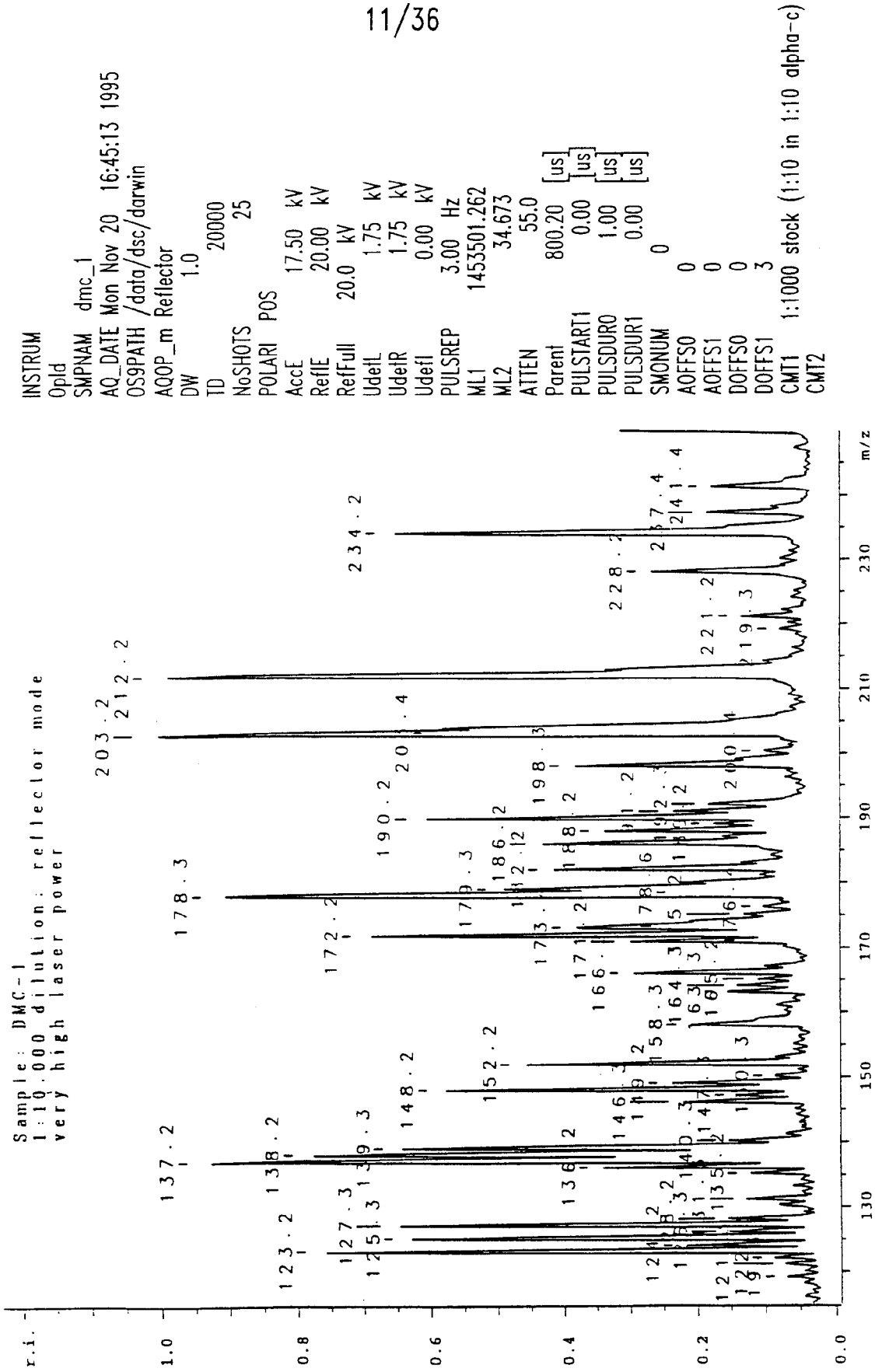


FIGURE 10

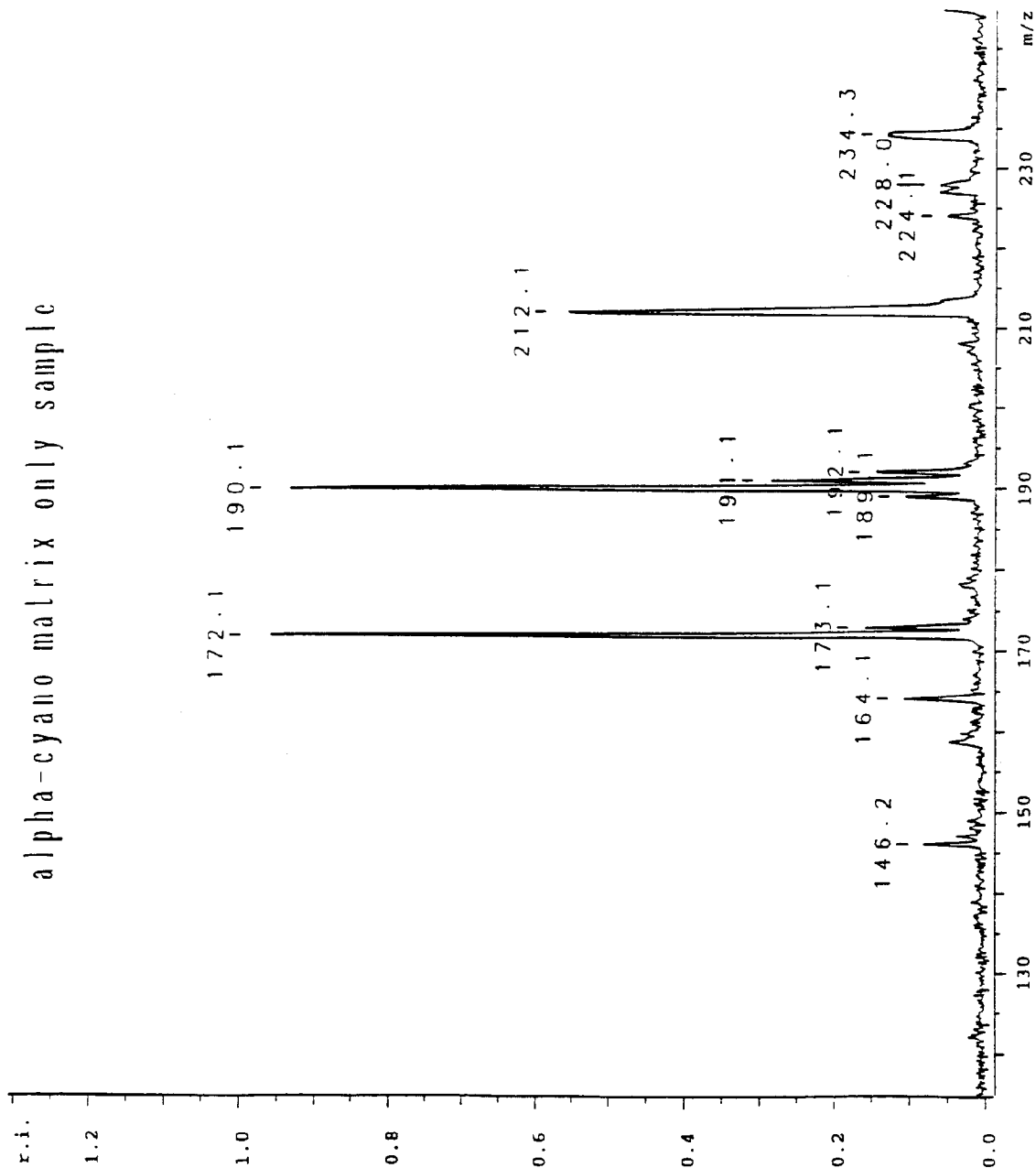
11/36

FIGURE 11



12/36

**FIGURE 12**  
alpha-cyano matrix only sample



INSTRUM  
OpId  
SMPNAM alpha\_c  
AQ\_DATE Mon Nov 20 14:04:49 1995  
OS9PATH /data/dsc/darwin  
AQOP\_m Reflector  
DW 4.0  
TD 10000  
NoSHOTS 25  
POLARI POS  
AccE 17.50 kV  
RefIE 20.00 kV  
RefFull 20.0 kV  
UdeIL 1.75 kV  
UdeIR 1.70 kV  
Udefl 0.00 kV  
PULSREP 3.00 Hz  
ML1 1454770.603  
ML2 37.845  
ATTEN 58.0  
Parent 800.20 [us]  
PULSTART1 0.00 [us]  
PULSDUR0 1.00 [us]  
PULSDUR1 0.00 [us]  
SMONUM 0  
AOFFS0 0  
AOFFS1 0  
DOFFS0 0  
DOFFS1 3  
CMT1 neat  
CMT2

13/36

Variable	Weight	Mass Spec	Photocleavable	5'-Aminohexyl Tailed
Weight	Range	Sensitivity	Linker	Oligonucleotide
Component	Adjuster	Enhancer		

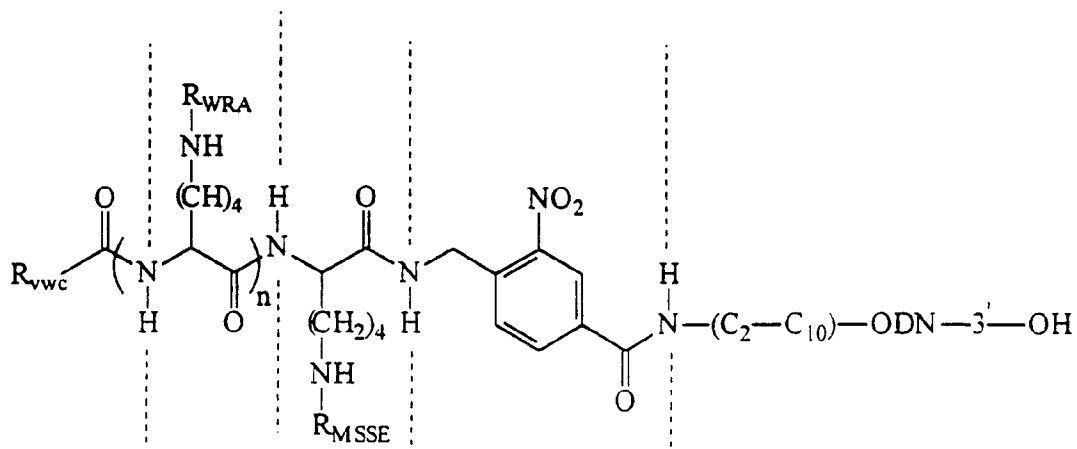
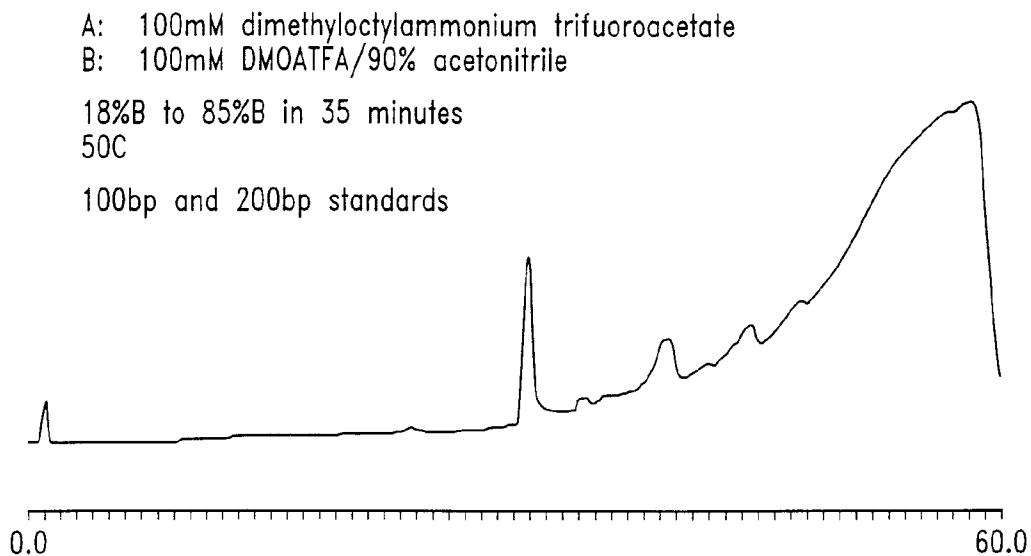
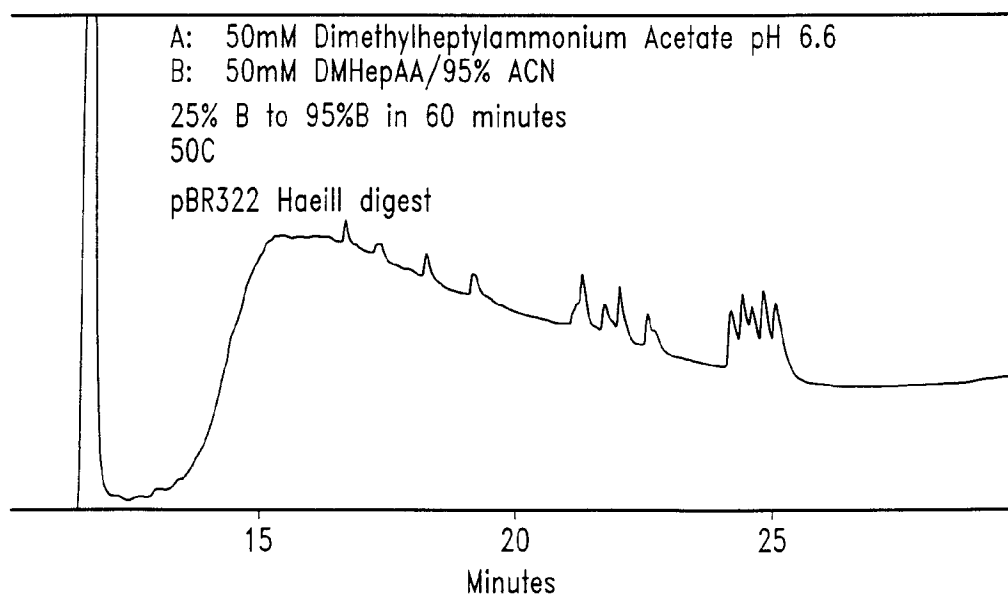


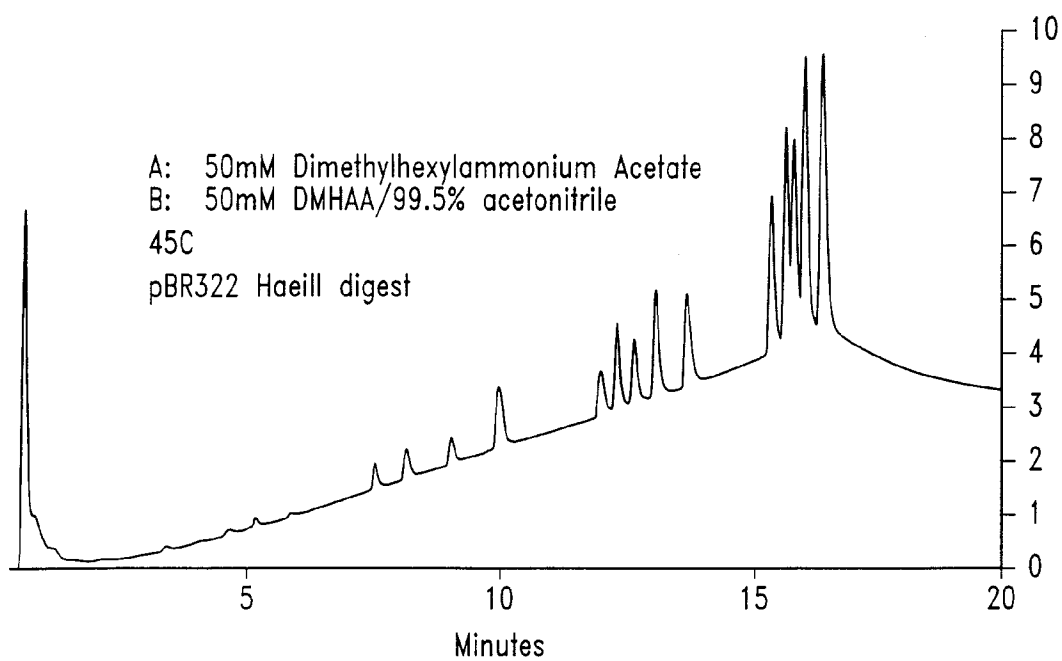
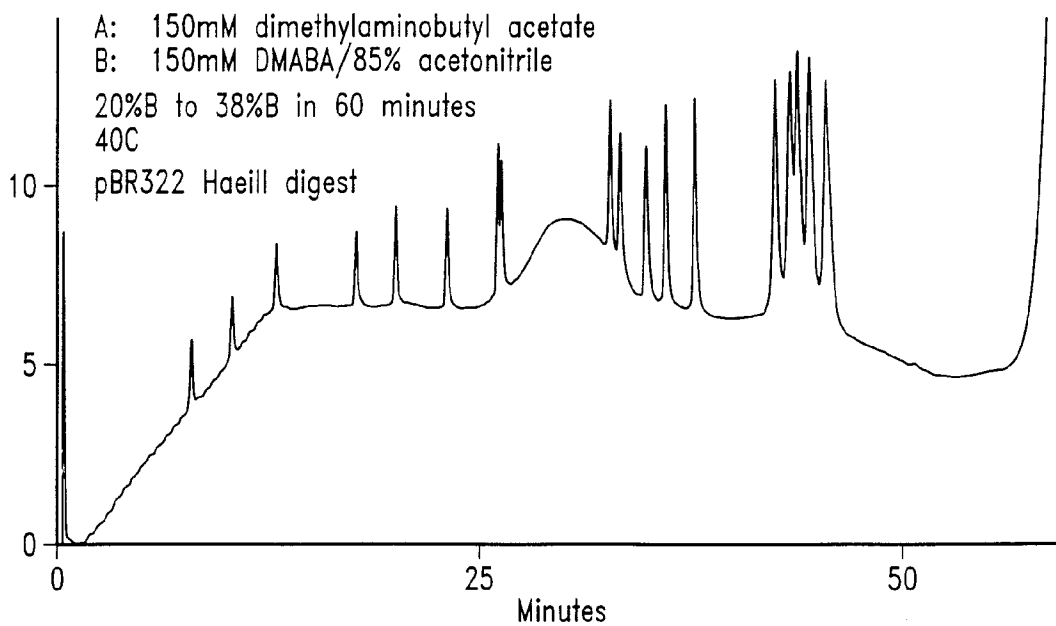
FIGURE 13



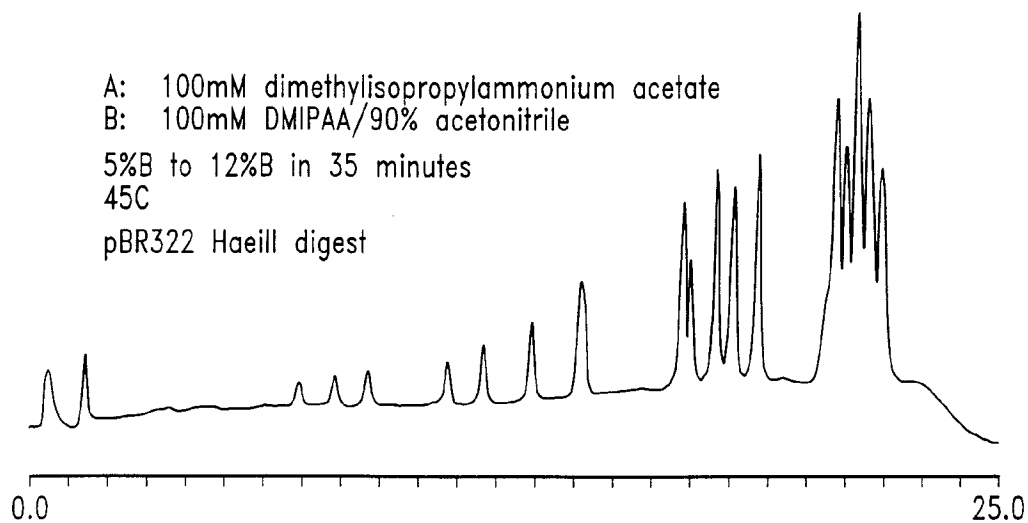
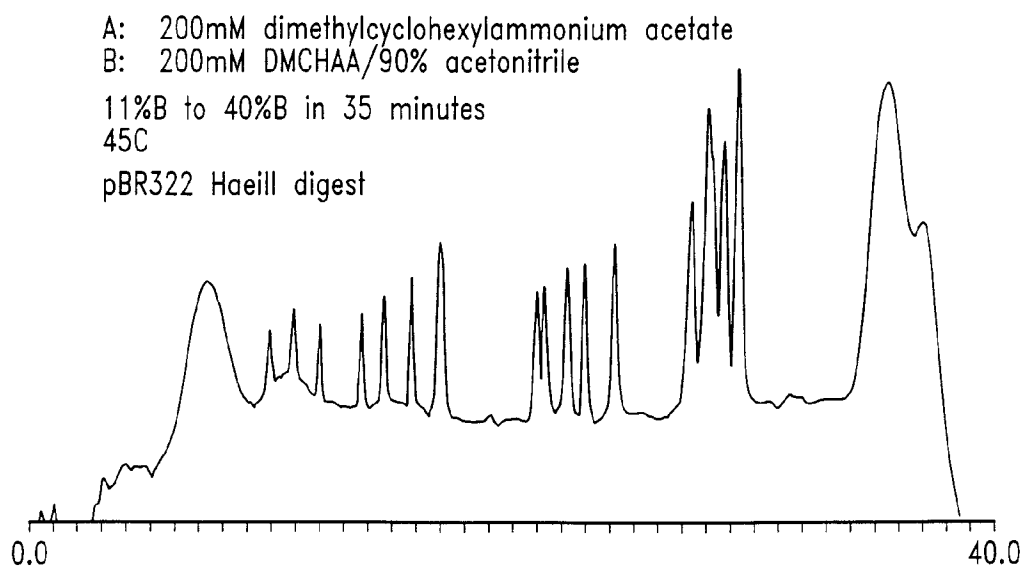
14/36

**FIGURE 14A****FIGURE 14B**

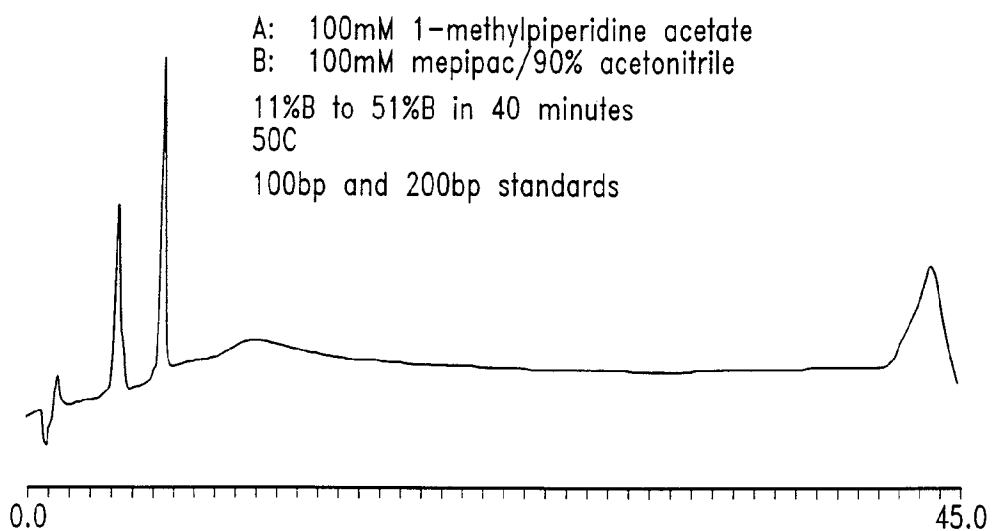
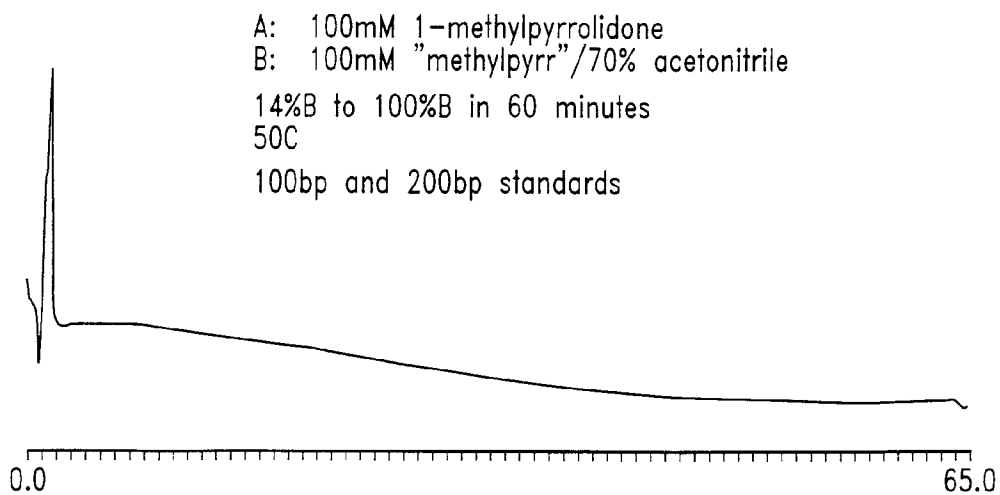
15/36

**FIGURE 14C****FIGURE 14D**

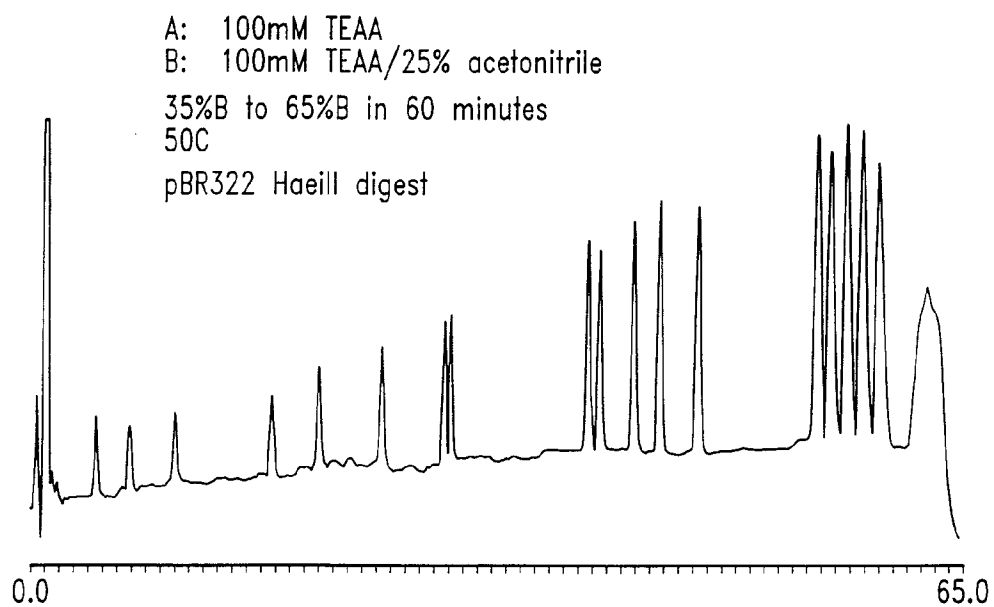
16/36

**FIGURE 14E****FIGURE 14F**

17/36

**FIGURE 14G****FIGURE 14H**

18/36

**FIGURE 14I**

# GENETIC FINGERPRINTING AND DIFFERENTIAL DISPLAY SYSTEMS

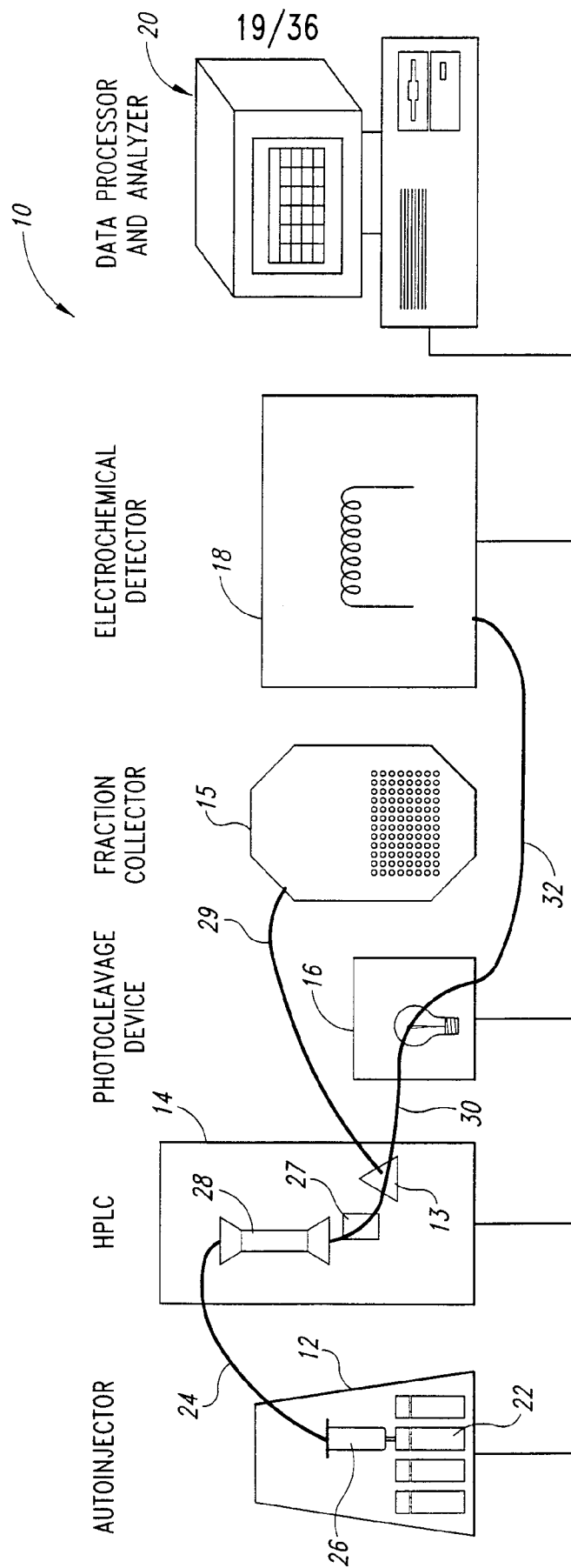


FIGURE 15

# GENETIC FINGERPRINTING AND DIFFERENTIAL DISPLAY SYSTEMS

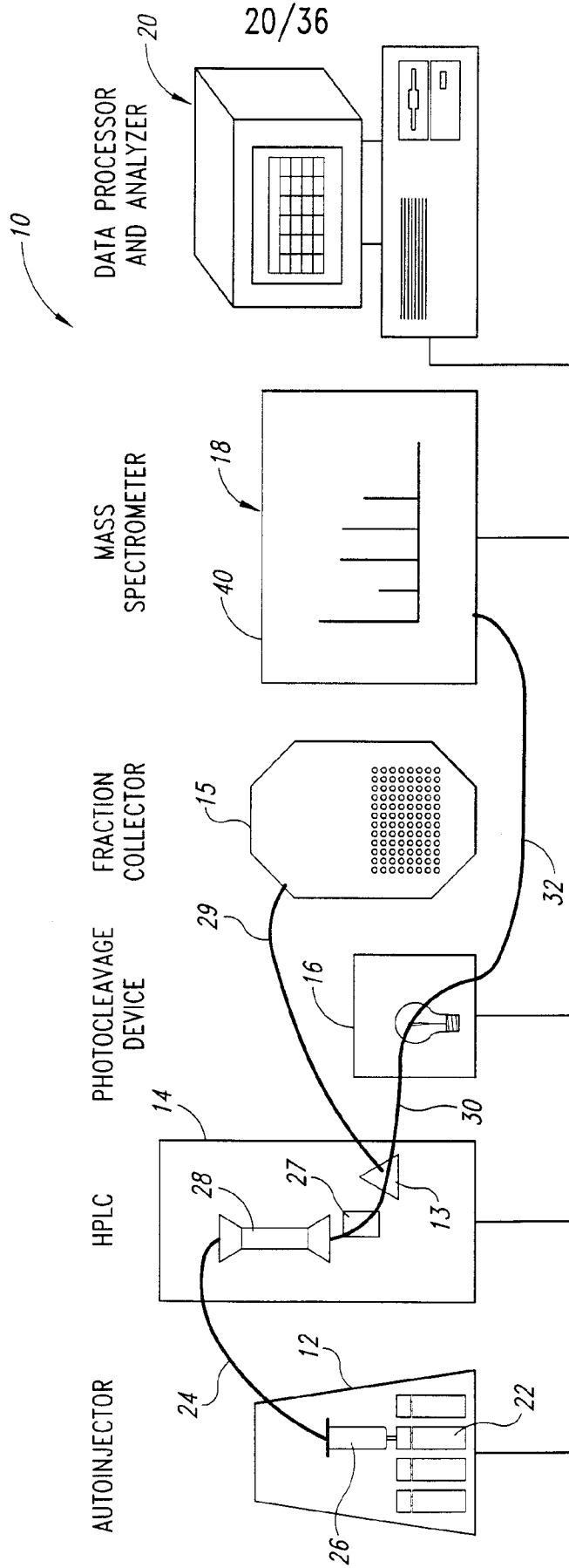
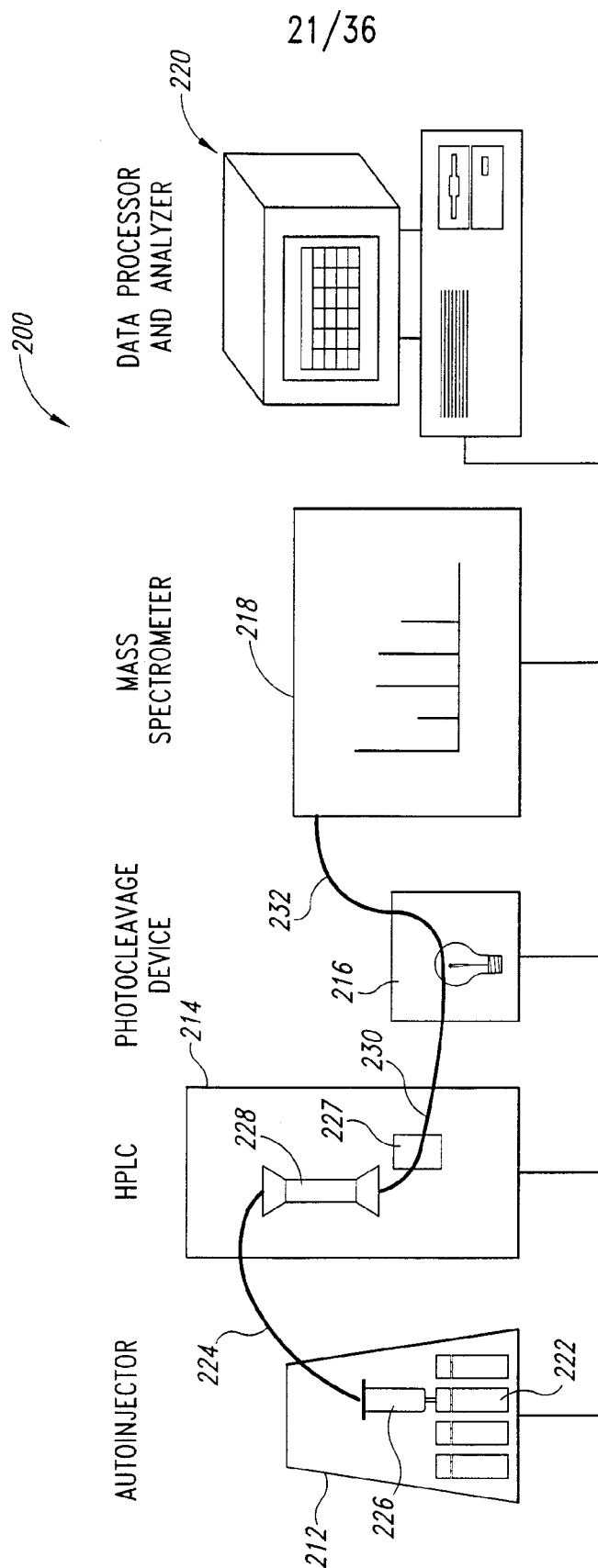


FIGURE 16

# ASSAY SYSTEMS



**FIGURE 17**



ASSAY SYSTEMS

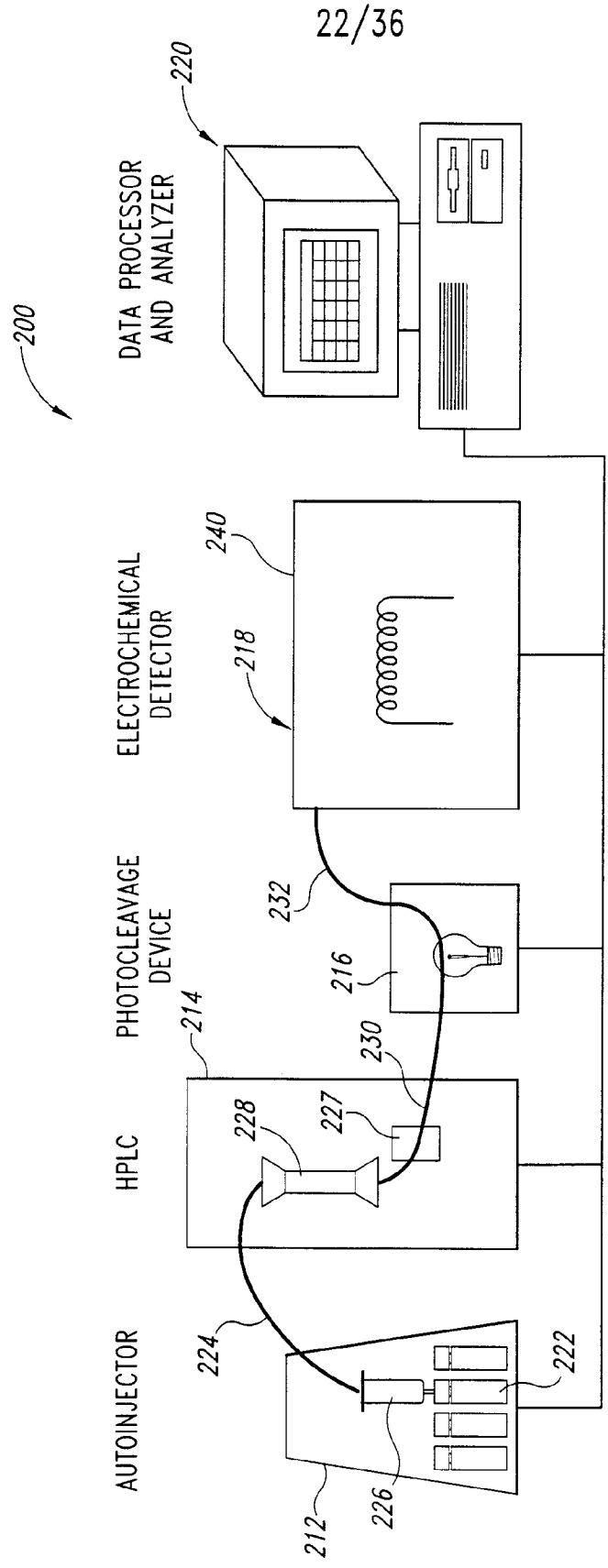
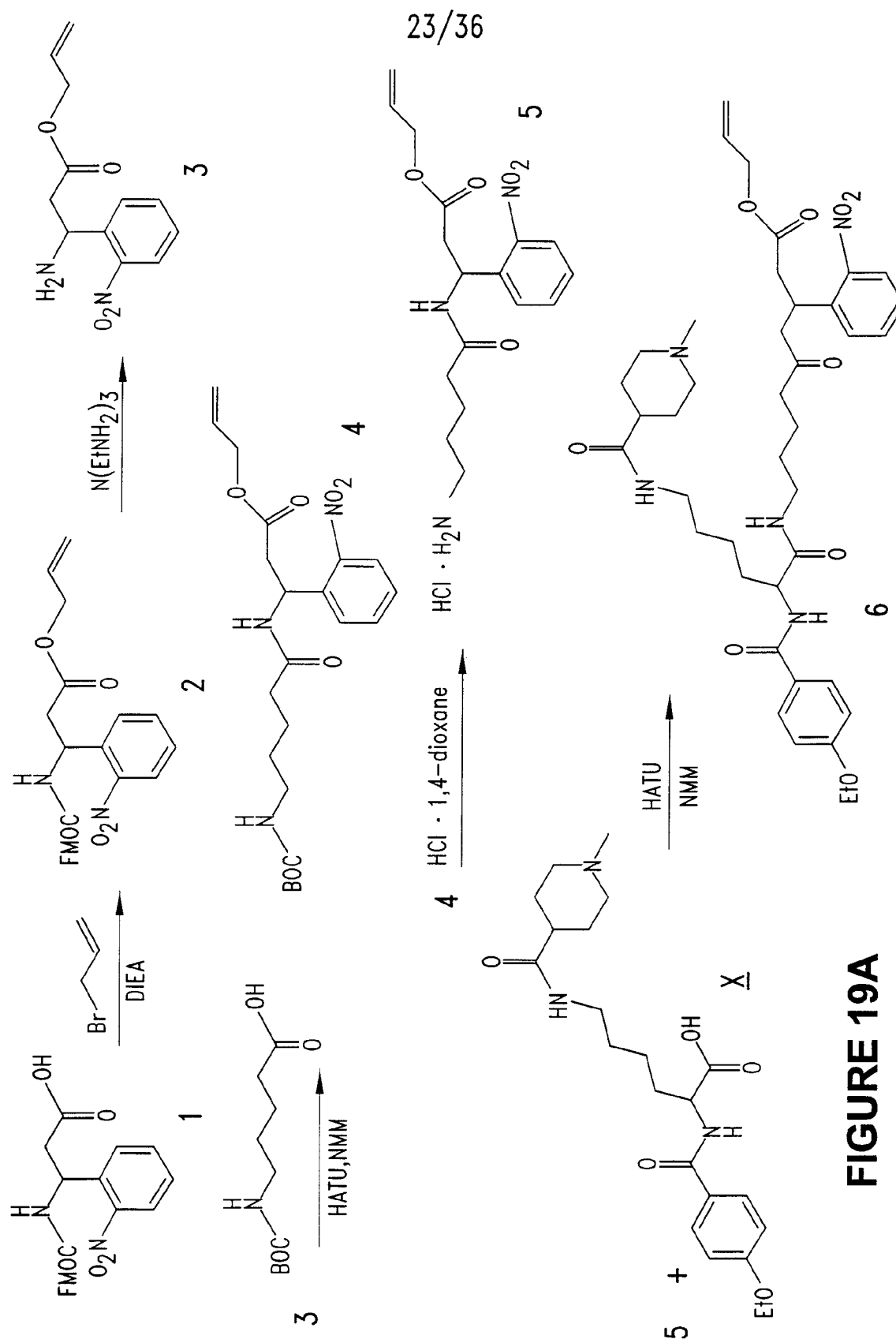


FIGURE 18



**FIGURE 19A**

24/36

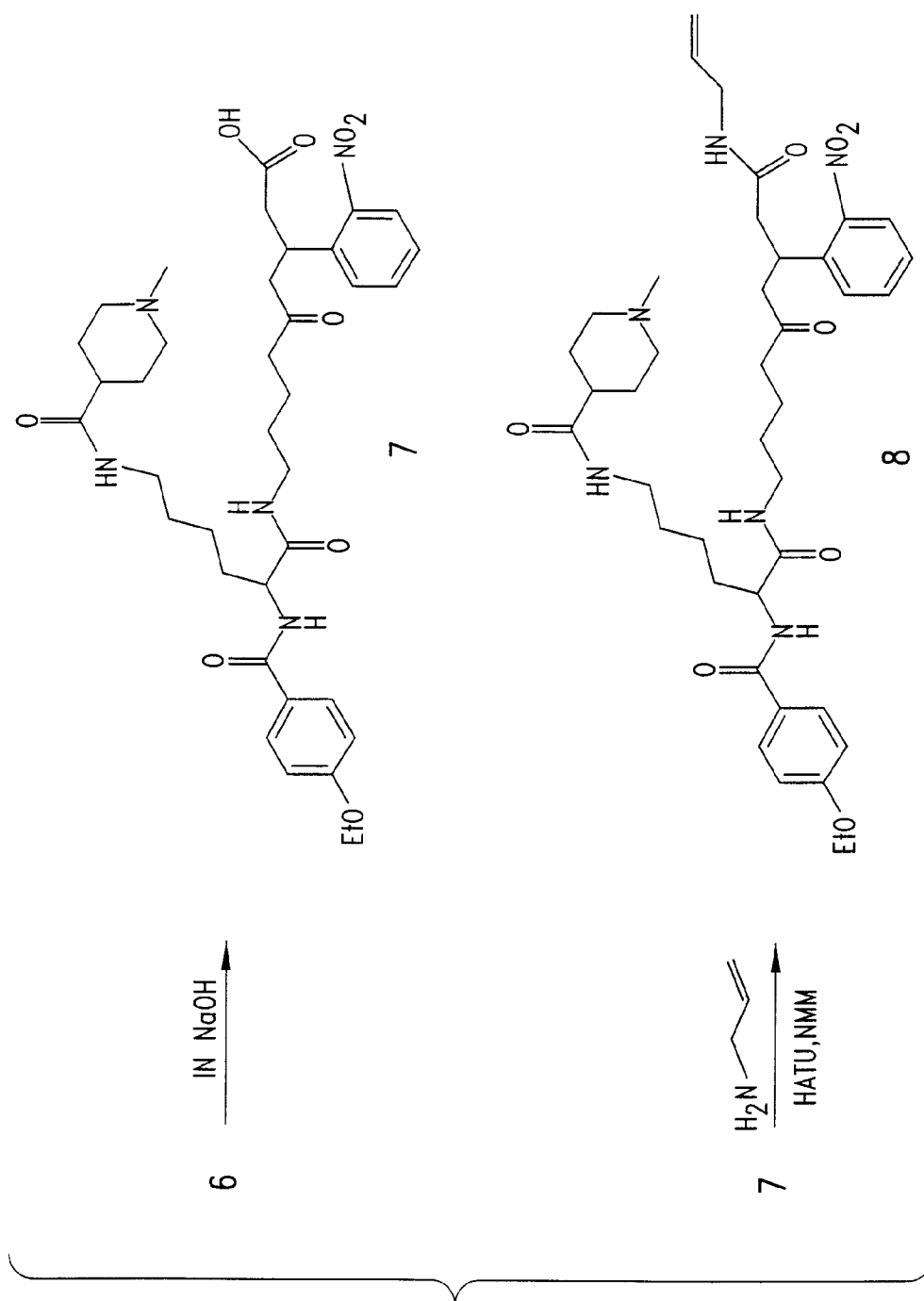


FIGURE 19B

25/36

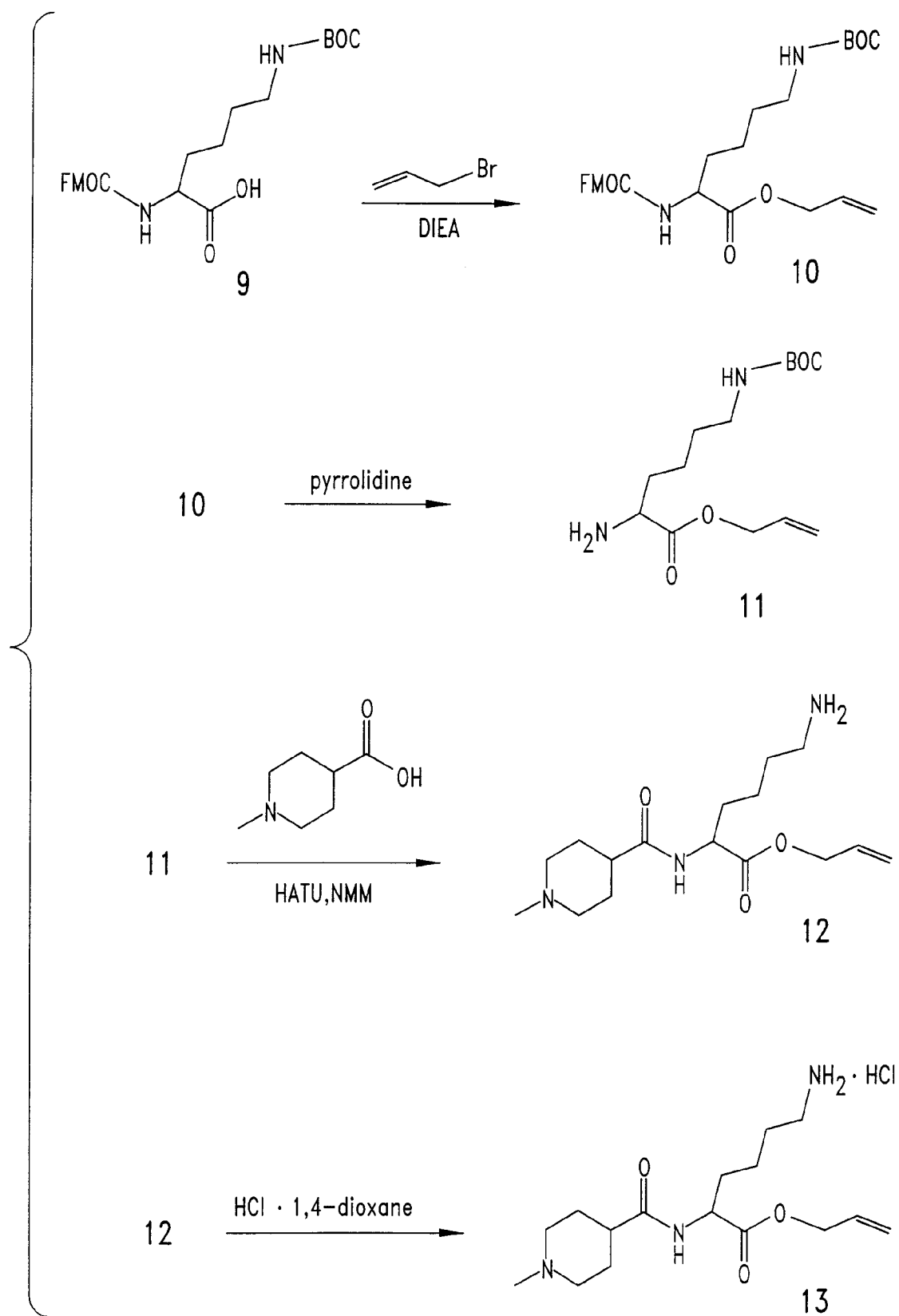
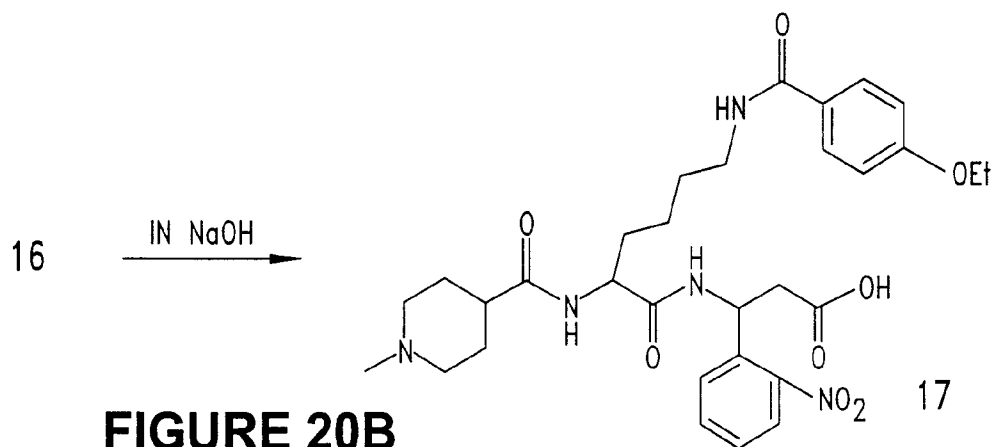
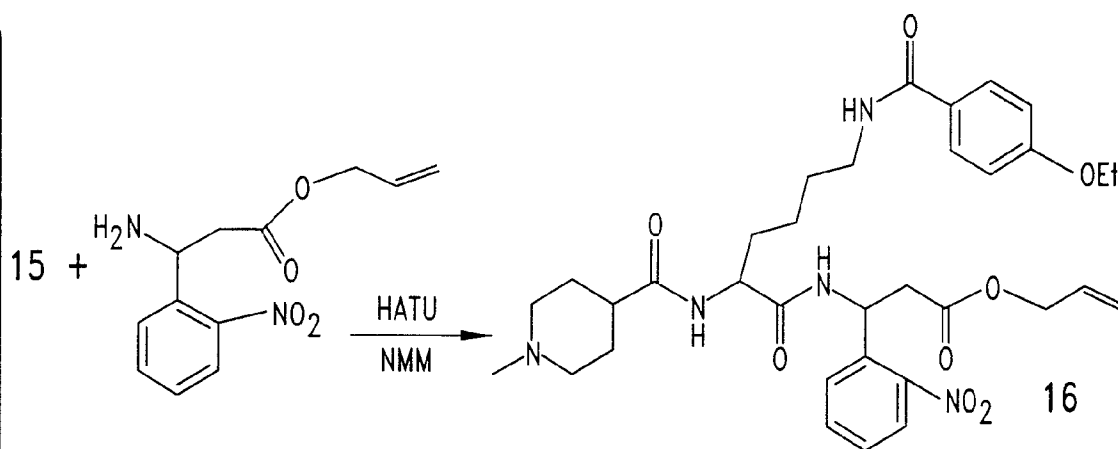
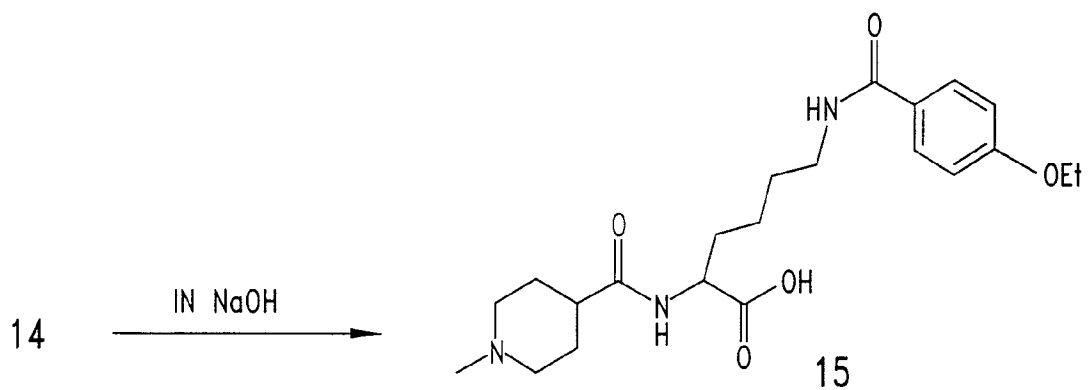
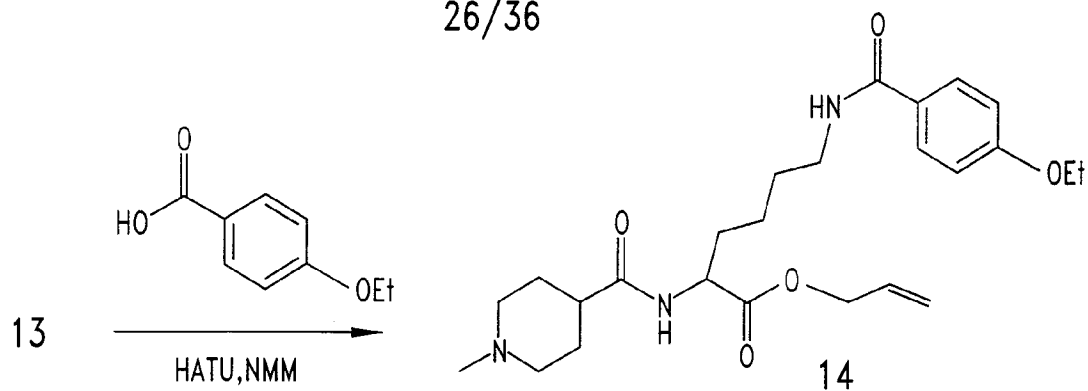
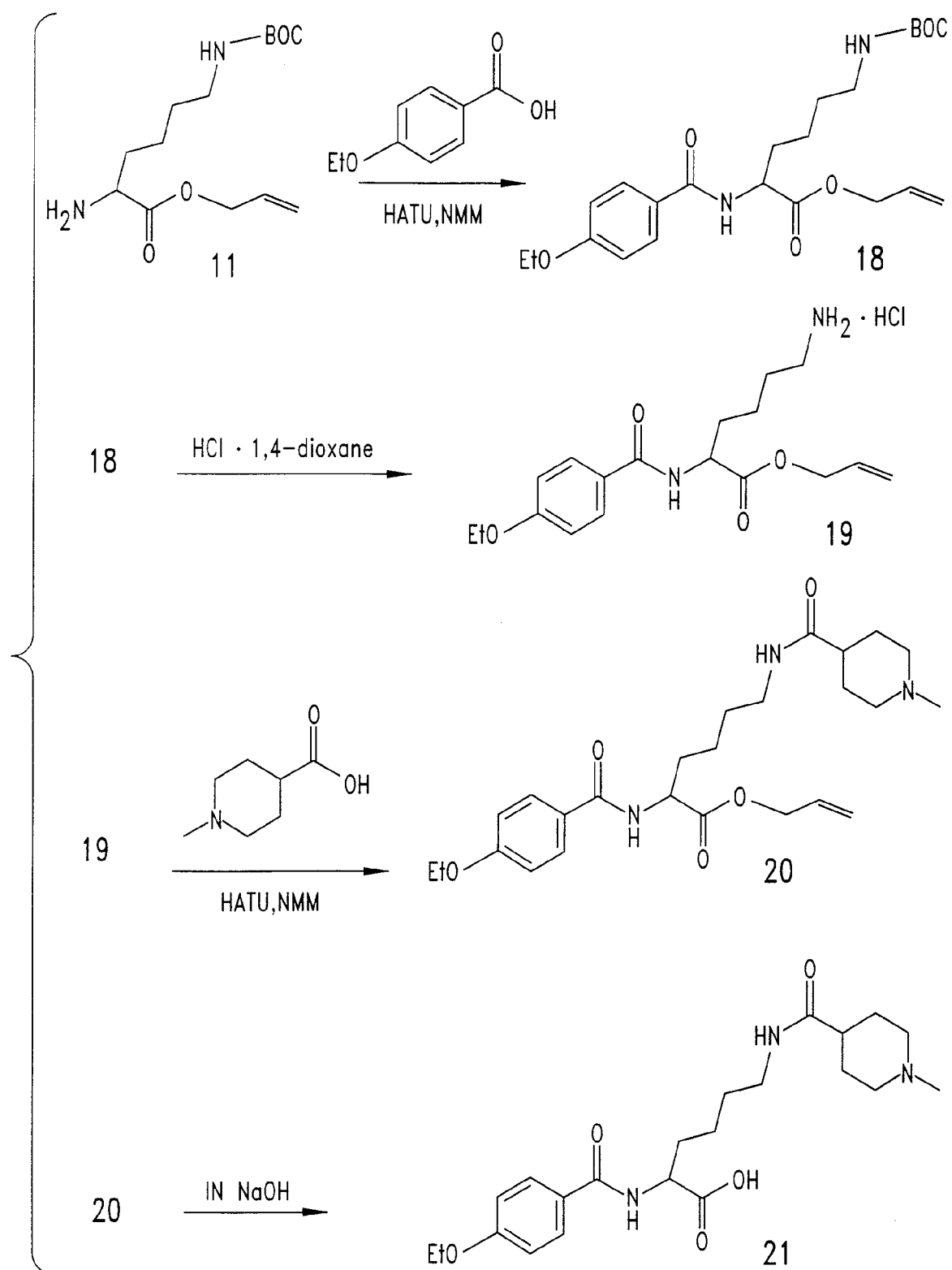


FIGURE 20A

26/36

**FIGURE 20B**

27/36

**FIGURE 21**

28/36

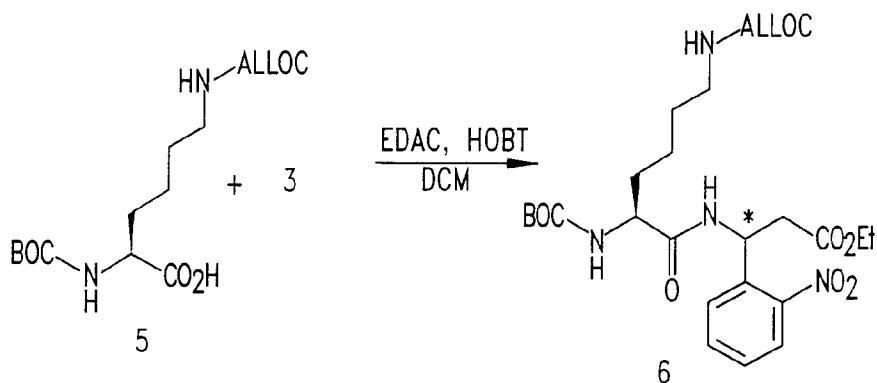
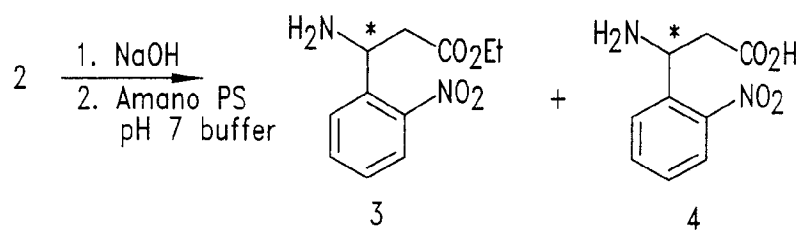
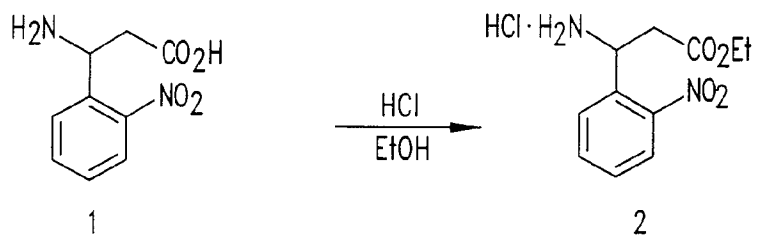


FIGURE 22A

29/36

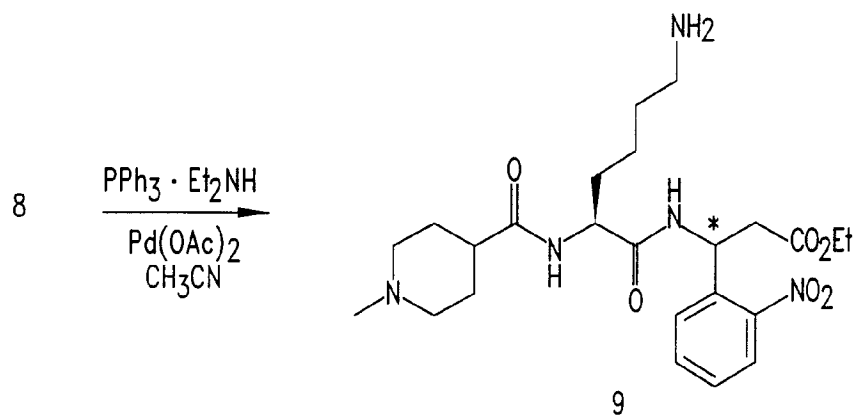
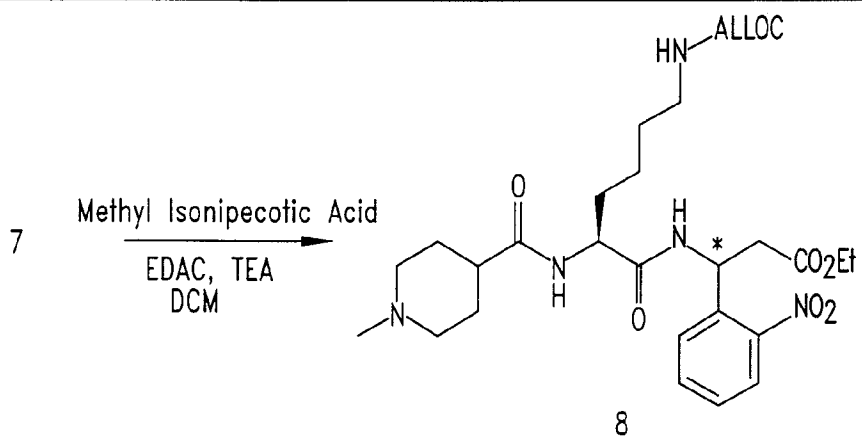
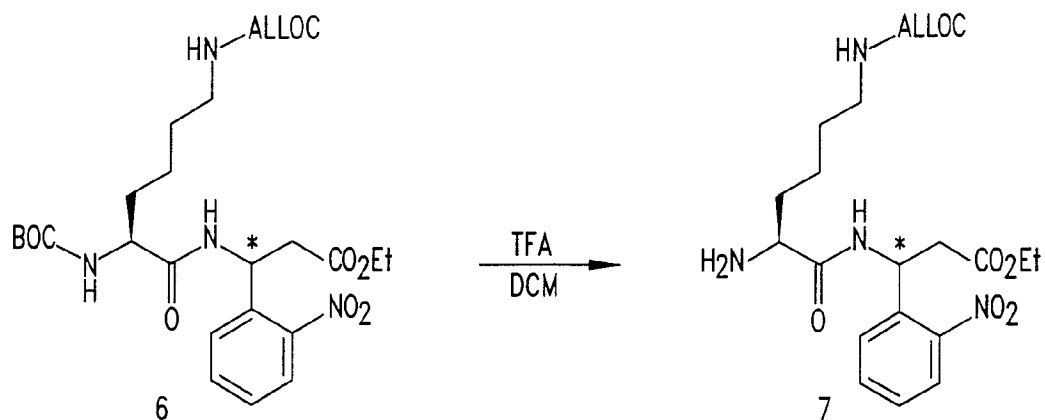


FIGURE 22B



30/36

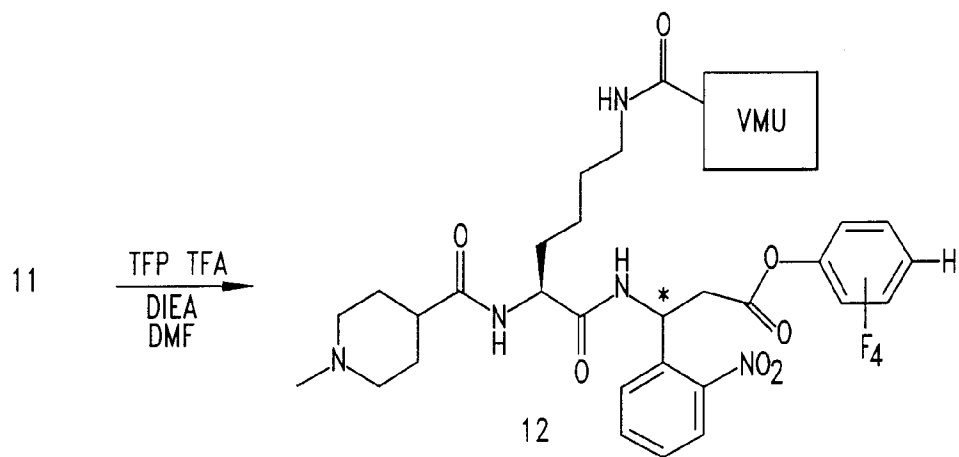
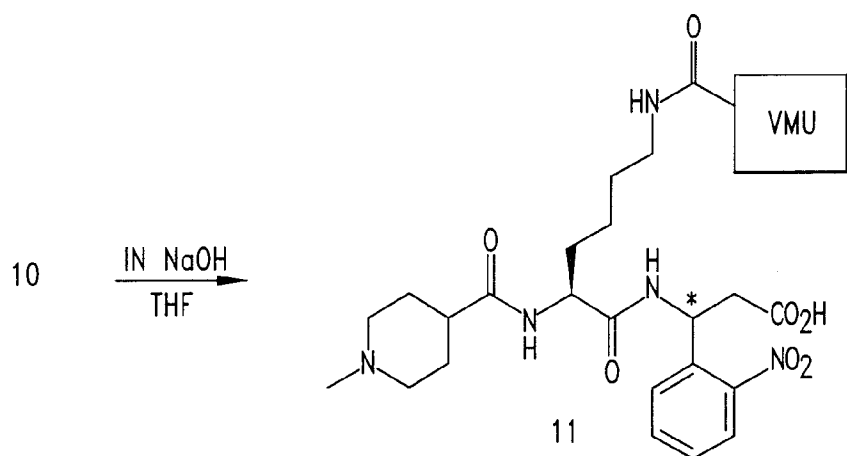
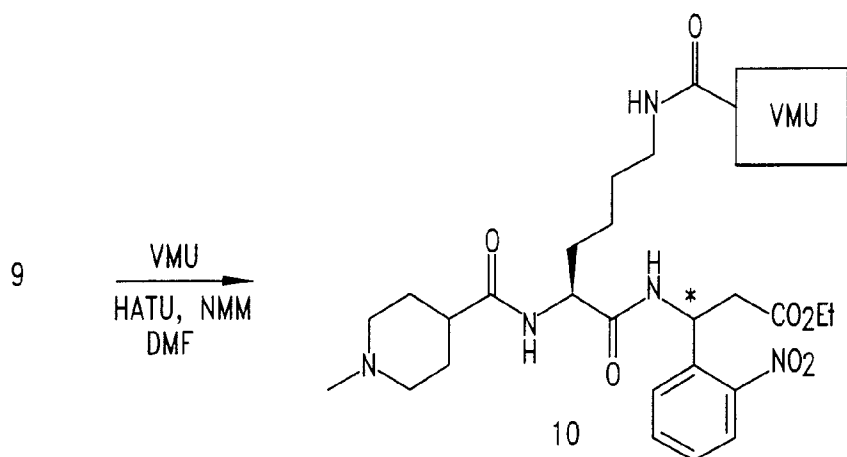
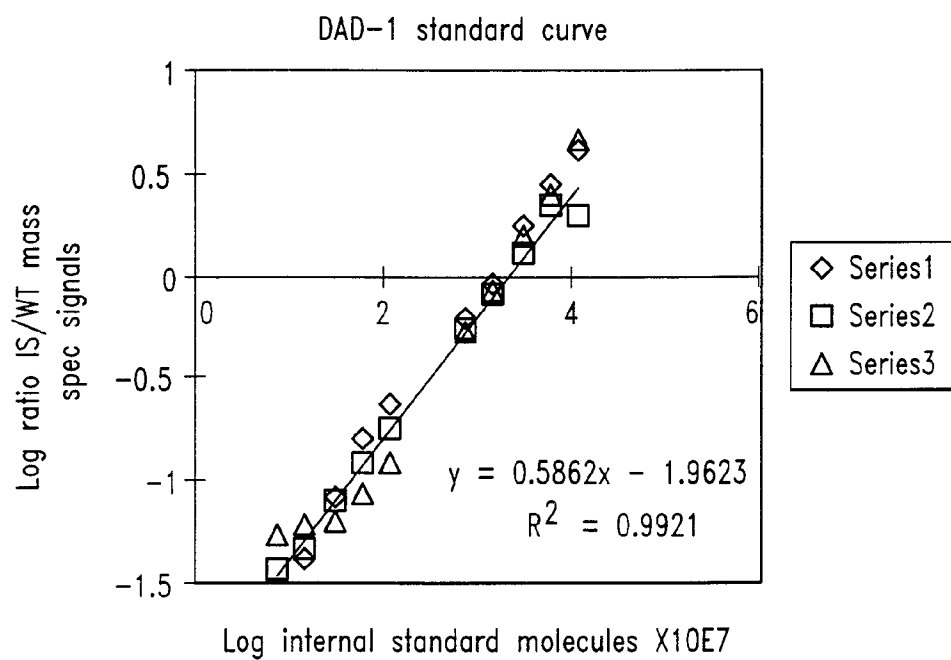
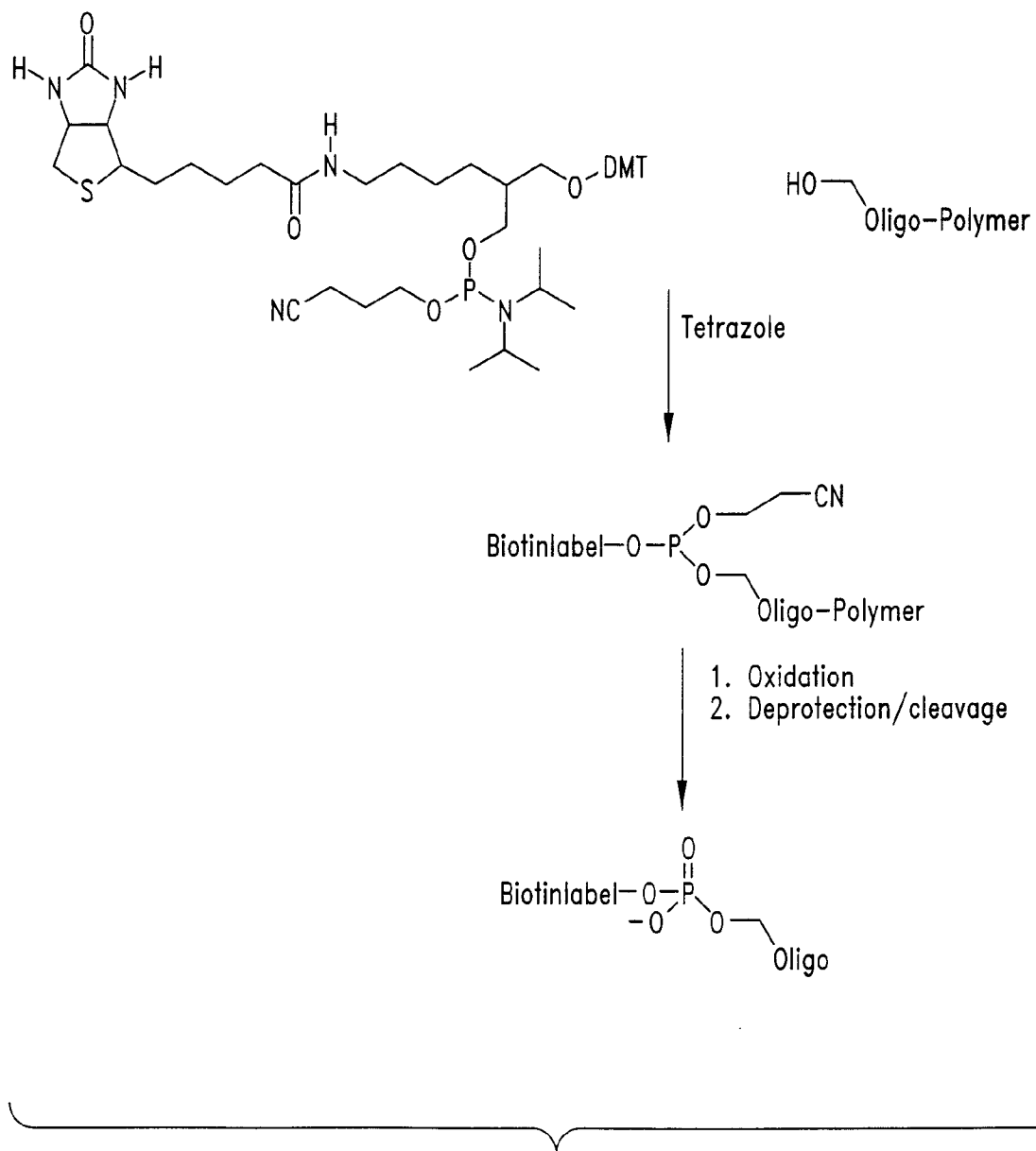


FIGURE 22C

31/36

**FIGURE 23**

32/36

**FIGURE 24**

33/36

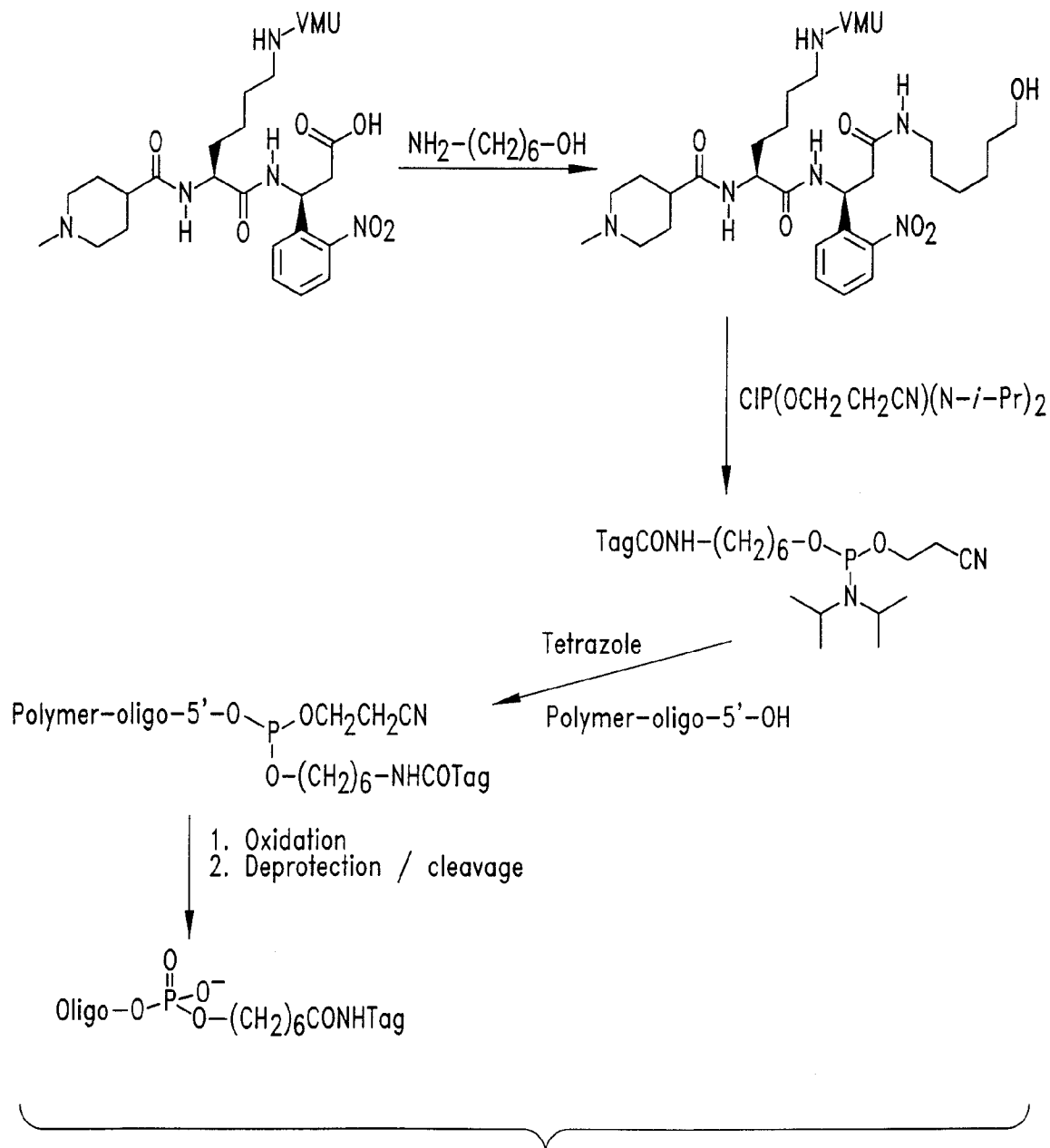
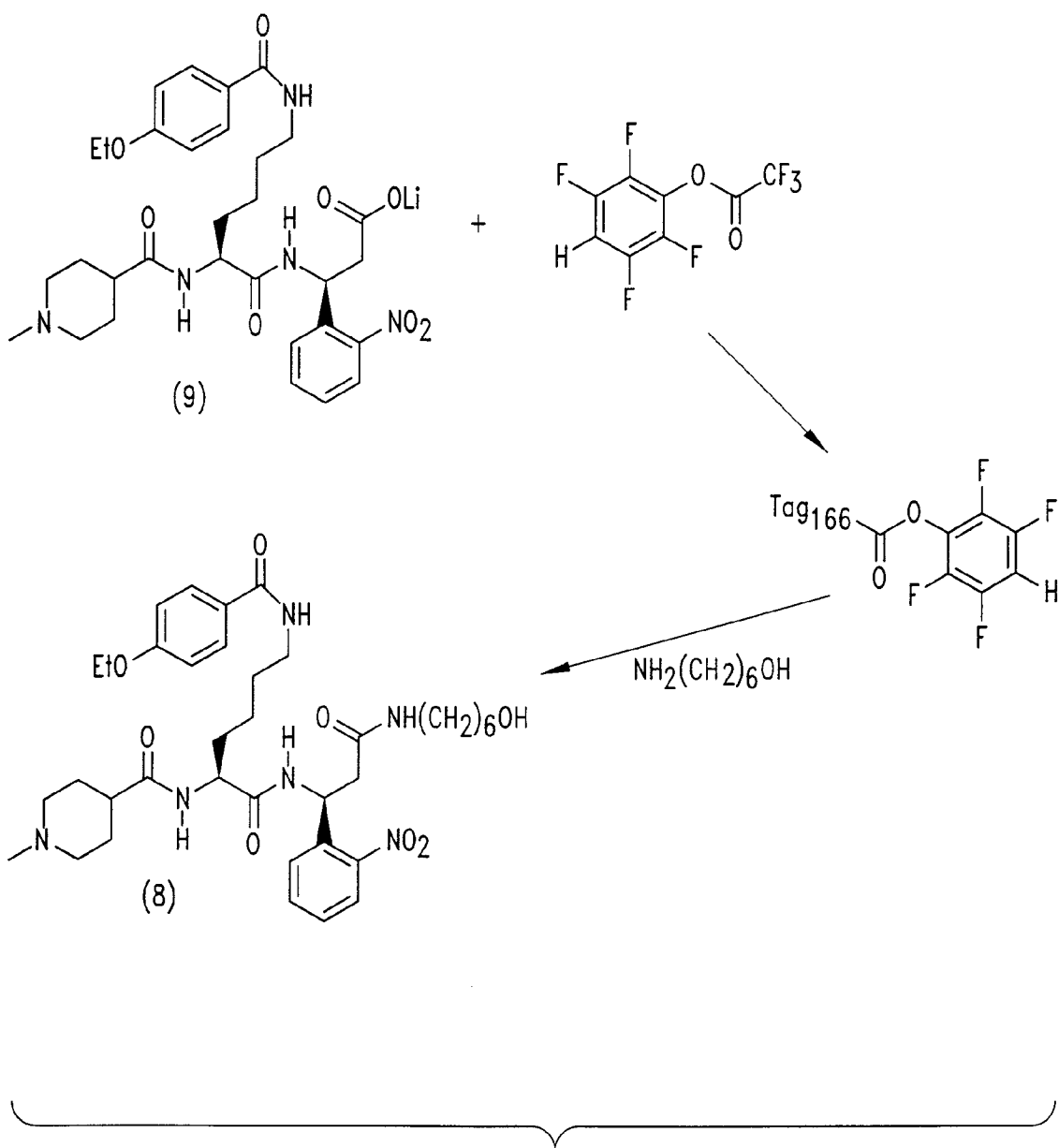


FIGURE 25

34/36

**FIGURE 26**

35/36

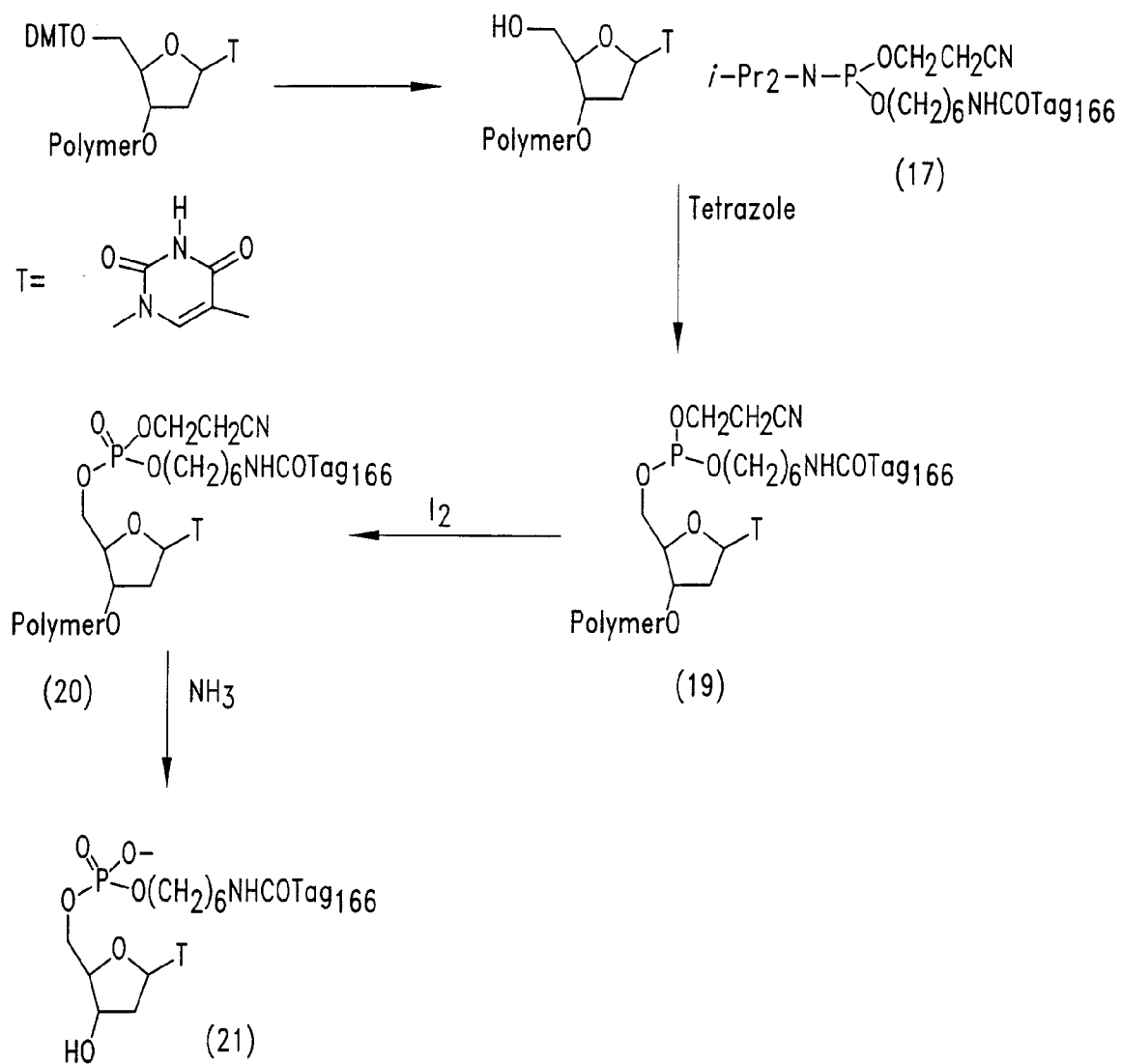
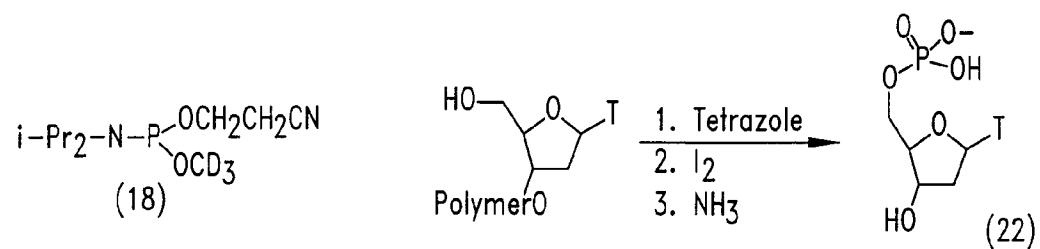


FIGURE 27

36/36

**FIGURE 28**

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/15008

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C07H21/00 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07H C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 118 605 A (URDEA MICHAEL S) 2 June 1992	1-3,5-7, 12,21-25
Y	see column 10, line 50 - column 11, line 44; claims 1-12	1-25
X	WO 95 04160 A (ISIS INNOVATION ;SOUTHERN EDWIN (GB); CUMMINS WILLIAM JONATHAN (GB) 9 February 1995	1,21-25
Y	see the whole document	1-25
X	WO 95 28640 A (UNIV COLUMBIA ;COLD SPRING HARBOR LAB (US); STILL W CLARK (US); WI) 26 October 1995 see claims 1,7,9,16-21,41,53	1
Y	WO 94 16101 A (KOESTER HUBERT) 21 July 1994 see claims 1-73	1-25
-/--		



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

22 February 1999

Date of mailing of the international search report

12/03/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Scott, J



## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/15008

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 95 25737 A (PENN STATE RES FOUND ;BENKOVIC STEPHEN J (US); WINOGRAD NICHOLAS () 28 September 1995 see the whole document ----	1-25
Y	JACOBSON K B ET AL: "APPLICATIONS OF MASS SPECTROMETRY TO DNA SEQUENCING" GENETIC ANALYSIS TECHNIQUES AND APPLICATIONS, vol. 8, no. 8, 1 December 1991, pages 223-229, XP000271820 see the whole document ----	1-25
Y	P.C.TOREN ET AL.: "Determination of Impurities in Nucleoside 3'-Phosphoramidites by Fast Atom Bombardment Mass Spectrometry" ANALYTICAL BIOCHEMISTRY, vol. 152, no. 2, 1986, pages 291-294, XP002094264 see the whole document ----	1-25
Y	B.B.BROWN ET AL.: "A Single-Bead Decode Strategy Using Electrospray Ionization Mass Spectrometry and a New Photolabile Linker : 3-Amino-3-(2-nitrophenyl)Propionic Acid." MOLECULAR DIVERSITY, vol. 1, 1995, pages 4-12, XP002094265 see the whole document ----	1-25
Y	WO 95 14108 A (AMERSHAM INT PLC ;REEVE MICHAEL ALAN (GB); HOWE ROLAND PAUL (GB); 26 May 1995 see abstract; claims 1-28 ----	1-25
P,X	WO 97 27331 A (DARWIN MOLECULAR CORP ;TABONE JOHN C (US); MULLIGAN JOHN T (US); H) 31 July 1997 see the whole document ----	1-25
P,X	WO 97 27325 A (DARWIN MOLECULAR CORP ;TABONE JOHN C (US); MULLIGAN JOHN T (US); H) 31 July 1997 see the whole document ----	1-25
P,X	WO 97 27327 A (DARWIN MOLECULAR CORP ;TABONE JOHN C (US); MULLIGAN JOHN T (US); H) 31 July 1997 see the whole document -----	1-25

# INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/US 98/15008

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5118605 A	02-06-1992	US 4775619 A	04-10-1988
		AT 133714 T	15-02-1996
		AT 168724 T	15-08-1998
		DE 3854969 D	14-03-1996
		DE 3854969 T	30-05-1996
		DE 3856224 D	27-08-1998
		DE 3856224 T	03-12-1998
		EP 0360940 A	04-04-1990
		EP 0703296 A	27-03-1996
		ES 2083955 T	01-05-1996
		JP 2092300 A	03-04-1990
		JP 2676535 B	17-11-1997
		US 5258506 A	02-11-1993
		US 5545730 A	13-08-1996
		US 5578717 A	26-11-1996
		US 5552538 A	03-09-1996
		US 5430136 A	04-07-1995
		US 5367066 A	22-11-1994
		US 5380833 A	10-01-1995
WO 9504160 A	09-02-1995	AT 159767 T	15-11-1997
		AU 695349 B	13-08-1998
		AU 7269194 A	28-02-1995
		CA 2168010 A	09-02-1995
		CN 1131440 A	18-09-1996
		DE 69406544 D	04-12-1997
		DE 69406544 T	26-02-1998
		DK 711362 T	22-12-1997
		EP 0711362 A	15-05-1996
		EP 0778280 A	11-06-1997
		ES 2108479 T	16-12-1997
		FI 960403 A	29-01-1996
		HU 73802 A	30-09-1996
		JP 9501830 T	25-02-1997
		NO 960370 A	28-03-1996
		US 5770367 A	23-06-1998
WO 9528640 A	26-10-1995	US 5565324 A	15-10-1996
		AU 2292695 A	10-11-1995
		CA 2187792 A	26-10-1995
		CN 1151793 A	11-06-1997
		EP 0755514 A	29-01-1997
		HU 74985 A	28-03-1997
		JP 10502614 T	10-03-1998
		NO 964332 A	03-12-1996
		US 5789172 A	04-08-1998
WO 9416101 A	21-07-1994	AU 694940 B	06-08-1998
		AU 5992994 A	15-08-1994
		AU 9137998 A	14-01-1999
		CA 2153387 A	21-07-1994
		EP 0679196 A	02-11-1995
		JP 8509857 T	22-10-1996
		US 5547835 A	20-08-1996
		US 5605798 A	25-02-1997
		US 5691141 A	25-11-1997
WO 9525737 A	28-09-1995	EP 0751950 A	08-01-1997

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/15008

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9525737 A		JP 9510711 T	28-10-1997
		US 5834195 A	10-11-1998
WO 9514108 A	26-05-1995	EP 0765401 A	02-04-1997
		JP 9505397 T	27-05-1997
		US 5849542 A	15-12-1998
WO 9727331 A	31-07-1997	AU 2247397 A	20-08-1997
		CA 2243560 A	31-07-1997
		CZ 9802184 A	16-12-1998
		EP 0868535 A	07-10-1998
WO 9727325 A	31-07-1997	AU 2244897 A	20-08-1997
		CA 2243546 A	31-07-1997
		CZ 9802285 A	16-12-1998
		EP 0840804 A	13-05-1998
		NZ 331043 A	28-01-1999
WO 9727327 A	31-07-1997	AU 1707997 A	20-08-1997
		CA 2243989 A	31-07-1997
		CZ 9802183 A	16-12-1998
		EP 0850320 A	01-07-1998